

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- (i) APPLICANT: Ill, Charles R. et al.
- (ii) TITLE OF INVENTION: NOVEL VECTORS AND GENES EXHIBITING
INCREASED EXPRESSION
- (iii) NUMBER OF SEQUENCES: 11
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: LAHIVE & COCKFIELD, LLP
 - (B) STREET: 28 STATE STREET
 - (C) CITY: BOSTON
 - (D) STATE: MASSACHUSETTS
 - (E) COUNTRY: US
 - (F) ZIP: 02109
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Floppy disk
 - (B) COMPUTER: IBM PC compatible
 - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 - (D) SOFTWARE: PatentIn Release #1.0, Version #1.25
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER:
 - (B) FILING DATE: 04 DECEMBER 1998
 - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: US 60/067,614
 - (B) FILING DATE: 05 DECEMBER 1997
 - (C) APPLICATION NUMBER: US 60/071,596
 - (D) FILING DATE: 16 JANUARY 1998
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: REMILLARD, JANE E.
 - (B) REGISTRATION NUMBER: 38,872
 - (C) REFERENCE/DOCKET NUMBER: TTI-180
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: (617)227-7400
 - (B) TELEFAX: (617)742-4214

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 4374 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (ix) FEATURE:
 - (A) NAME/KEY: CDS



(B) LOCATION: 1..4374

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ATG GAA ATA GAG CTC TCC ACC TGC TTC TTT CTG TGC CTT TTG CGA TTC	48
Met Glu Ile Glu Leu Ser Thr Cys Phe Phe Leu Cys Leu Leu Arg Phe	
1 5 10 15	
TGC TTT AGT GCC ACC AGA AGA TAC TAC CTG GGT GCA GTG GAA CTG TCA	96
Cys Phe Ser Ala Thr Arg Arg Tyr Tyr Leu Gly Ala Val Glu Leu Ser	
20 25 30	
TGG GAC TAT ATG CAA AGT GAT CTC GGA GAG CTG CCT GTG GAC GCA AGA	144
Trp Asp Tyr Met Gln Ser Asp Leu Gly Glu Leu Pro Val Asp Ala Arg	
35 40 45	
TTT CCT CCT CGC GTG CCA AAA TCT TTT CCA TTC AAC ACC TCA GTC GTG	192
Phe Pro Pro Arg Val Pro Lys Ser Phe Pro Phe Asn Thr Ser Val Val	
50 55 60	
TAC AAA AAG ACT CTG TTT GTA GAA TTC ACG GTT CAC CTT TTC AAC ATC	240
Tyr Lys Lys Thr Leu Phe Val Glu Phe Thr Val His Leu Phe Asn Ile	
65 70 75 80	
GCT AAG CCA AGG CCA CCC TGG ATG GGT CTG CTA GGT CCT ACC ATC CAA	288
Ala Lys Pro Arg Pro Pro Trp Met Gly Leu Leu Gly Pro Thr Ile Gln	
85 90 95	
GCT GAG GTT TAT GAT ACA GTG GTC ATT ACA CTT AAG AAC ATG GCT TCC	336
Ala Glu Val Tyr Asp Thr Val Val Ile Thr Leu Lys Asn Met Ala Ser	
100 105 110	
CAT CCT GTC TCC CTT CAT GCT GTT GGT GTA TCC TAC TGG AAA GCT TCT	384
His Pro Val Ser Leu His Ala Val Gly Val Ser Tyr Trp Lys Ala Ser	
115 120 125	
GAG GGA GCT GAA TAT GAT GAT CAG ACC AGT CAA AGG GAG AAA GAA GAT	432
Glu Gly Ala Glu Tyr Asp Asp Gln Thr Ser Gln Arg Glu Lys Glu Asp	
130 135 140	
GAT AAA GTC TTC CCT GGT GGA AGC CAT ACA TAT GTC TGG CAA GTC CTG	480
Asp Lys Val Phe Pro Gly Gly Ser His Thr Tyr Val Trp Gln Val Leu	
145 150 155 160	
AAA GAG AAT GGT CCA ATG GCC TCC GAC CCA CTG TGC CTT ACC TAC TCA	528
Lys Glu Asn Gly Pro Met Ala Ser Asp Pro Leu Cys Leu Thr Tyr Ser	
165 170 175	
TAT CTT TCT CAT GTG GAC CTG GTT AAA GAC TTG AAT TCA GGC CTC ATT	576
Tyr Leu Ser His Val Asp Leu Val Lys Asp Leu Asn Ser Gly Leu Ile	
180 185 190	
GGA GCC CTA CTA GTA TGT AGA GAA GGG AGT CTG GCC AAG GAA AAG ACA	624
Gly Ala Leu Leu Val Cys Arg Glu Gly Ser Leu Ala Lys Glu Lys Thr	
195 200 205	
CAG ACC TTG CAC AAA TTT ATA CTA CTT TTT GCT GTA TTT GAT GAA GGG	672

Gln	Thr	Leu	His	Lys	Phe	Ile	Leu	Leu	Phe	Ala	Val	Phe	Asp	Glu	Gly	
210						215					220					
AAA	AGT	TGG	CAC	TCA	GAA	ACA	AAG	AAC	TCC	CTC	ATG	CAA	GAT	AGG	GAT	720
Lys	Ser	Trp	His	Ser	Glu	Thr	Lys	Asn	Ser	Leu	Met	Gln	Asp	Arg	Asp	
225					230					235					240	
GCT	GCA	TCT	GCT	CGG	GCC	TGG	CCT	AAA	ATG	CAC	ACA	GTC	AAT	GGT	TAT	768
Ala	Ala	Ser	Ala	Arg	Ala	Trp	Pro	Lys	Met	His	Thr	Val	Asn	Gly	Tyr	
				245					250					255		
GTA	AAC	AGG	AGC	CTG	CCA	GGA	CTG	ATT	GGA	TGC	CAC	AGG	AAA	TCA	GTC	816
Val	Asn	Arg	Ser	Leu	Pro	Gly	Leu	Ile	Gly	Cys	His	Arg	Lys	Ser	Val	
			260					265					270			
TAT	TGG	CAT	GTT	ATA	GGA	ATG	GGC	ACC	ACT	CCT	GAA	GTG	CAC	TCA	ATA	864
Tyr	Trp	His	Val	Ile	Gly	Met	Gly	Thr	Thr	Pro	Glu	Val	His	Ser	Ile	
	275						280					285				
TTC	CTC	GAA	GGA	CAC	ACA	TTT	CTT	GTT	AGA	AAC	CAT	CGC	CAG	GCG	TCC	912
Phe	Leu	Glu	Gly	His	Thr	Phe	Leu	Val	Arg	Asn	His	Arg	Gln	Ala	Ser	
	290					295					300					
TTG	GAA	ATC	TCG	CCA	ATA	ACT	TTC	CTT	ACT	GCT	CAA	ACA	CTC	CTC	ATG	960
Leu	Glu	Ile	Ser	Pro	Ile	Thr	Phe	Leu	Thr	Ala	Gln	Thr	Leu	Leu	Met	
305					310					315					320	
GAC	CTT	GGA	CAG	TTT	CTA	CTG	TTT	TGT	CAT	ATC	TCT	TCC	CAC	CAA	CAT	1008
Asp	Leu	Gly	Gln	Phe	Leu	Leu	Phe	Cys	His	Ile	Ser	Ser	His	Gln	His	
				325					330					335		
GAT	GGC	ATG	GAA	GCT	TAT	GTC	AAA	GTA	GAC	AGC	TGT	CCA	GAG	GAA	CCC	1056
Asp	Gly	Met	Glu	Ala	Tyr	Val	Lys	Val	Asp	Ser	Cys	Pro	Glu	Glu	Pro	
			340					345					350			
CAA	CTA	CGA	ATG	AAA	AAT	AAT	GAA	GAA	GCG	GAA	GAC	TAT	GAT	GAT	GAT	1104
Gln	Leu	Arg	Met	Lys	Asn	Asn	Glu	Glu	Ala	Glu	Asp	Tyr	Asp	Asp	Asp	
		355					360					365				
CTT	ACC	GAT	TCT	GAA	ATG	GAT	GTG	GTC	AGA	TTT	GAT	GAT	GAC	AAC	TCT	1152
Leu	Thr	Asp	Ser	Glu	Met	Asp	Val	Val	Arg	Phe	Asp	Asp	Asp	Asn	Ser	
	370					375					380					
CCT	TCC	TTT	ATC	CAA	ATT	CGC	TCA	GTT	GCC	AAG	AAG	CAT	CCT	AAA	ACT	1200
Pro	Ser	Phe	Ile	Gln	Ile	Arg	Ser	Val	Ala	Lys	Lys	His	Pro	Lys	Thr	
385					390					395					400	
TGG	GTA	CAT	TAC	ATT	GCT	GCT	GAA	GAG	GAG	GAC	TGG	GAC	TAT	GCT	CCC	1248
Trp	Val	His	Tyr	Ile	Ala	Ala	Glu	Glu	Glu	Asp	Trp	Asp	Tyr	Ala	Pro	
				405				410						415		
TTA	GTC	CTC	GCC	CCC	GAT	GAC	AGA	AGT	TAT	AAA	AGT	CAA	TAT	TTG	AAC	1296
Leu	Val	Leu	Ala	Pro	Asp	Asp	Arg	Ser	Tyr	Lys	Ser	Gln	Tyr	Leu	Asn	
			420					425					430			
AAT	GGC	CCT	CAG	CGG	ATT	GGA	AGG	AAG	TAC	AAA	AAA	GTC	CGA	TTT	ATG	1344
Asn	Gly	Pro	Gln	Arg	Ile	Gly	Arg	Lys	Tyr	Lys	Lys	Val	Arg	Phe	Met	

435						440						445						
GCA	TAC	ACA	GAT	GAA	ACC	TTT	AAG	ACT	CGT	GAA	GCT	ATT	CAG	CAT	GAA	1392		
Ala	Tyr	Thr	Asp	Glu	Thr	Phe	Lys	Thr	Arg	Glu	Ala	Ile	Gln	His	Glu			
450			455			460												
TCA	GGA	ATC	TTG	GGA	CCT	TTA	CTT	TAT	GGG	GAA	GTT	GGA	GAC	ACA	CTG	1440		
Ser	Gly	Ile	Leu	Gly	Pro	Leu	Leu	Tyr	Gly	Glu	Val	Gly	Asp	Thr	Leu			
465			470			475			480									
CTC	ATT	ATA	TTT	AAG	AAT	CAA	GCA	AGC	AGA	CCA	TAT	AAC	ATC	TAC	CCT	1488		
Leu	Ile	Ile	Phe	Lys	Asn	Gln	Ala	Ser	Arg	Pro	Tyr	Asn	Ile	Tyr	Pro			
485			490			495												
CAC	GGA	ATC	ACC	GAT	GTC	CGT	CCT	TTG	TAT	TCA	CGC	AGA	TTA	CCA	AAA	1536		
His	Gly	Ile	Thr	Asp	Val	Arg	Pro	Leu	Tyr	Ser	Arg	Arg	Leu	Pro	Lys			
500			505			510												
GGA	GTA	AAA	CAT	TTG	AAG	GAT	TTT	CCA	ATT	CTG	CCC	GGA	GAA	ATA	TTC	1584		
Gly	Val	Lys	His	Leu	Lys	Asp	Phe	Pro	Ile	Leu	Pro	Gly	Glu	Ile	Phe			
515			520			525												
AAA	TAT	AAA	TGG	ACA	GTG	ACT	GTA	GAA	GAT	GGG	CCA	ACT	AAA	TCA	GAT	1632		
Lys	Tyr	Lys	Trp	Thr	Val	Thr	Val	Glu	Asp	Gly	Pro	Thr	Lys	Ser	Asp			
530			535			540												
CCT	CGG	TGC	CTG	ACC	CGC	TAT	TAC	TCT	AGT	TTC	GTC	AAT	ATG	GAG	AGA	1680		
Pro	Arg	Cys	Leu	Thr	Arg	Tyr	Tyr	Ser	Ser	Phe	Val	Asn	Met	Glu	Arg			
545			550			555			560									
GAT	CTA	GCT	TCA	GGA	CTC	ATT	GGC	CCT	CTC	CTC	ATC	TGC	TAC	AAA	GAA	1728		
Asp	Leu	Ala	Ser	Gly	Leu	Ile	Gly	Pro	Leu	Leu	Ile	Cys	Tyr	Lys	Glu			
565			570			575												
TCT	GTA	GAT	CAA	AGA	GGA	AAC	CAG	ATA	ATG	TCA	GAC	AAG	AGG	AAT	GTC	1776		
Ser	Val	Asp	Gln	Arg	Gly	Asn	Gln	Ile	Met	Ser	Asp	Lys	Arg	Asn	Val			
580			585			590												
ATC	CTG	TTT	TCT	GTA	TTT	GAT	GAG	AAC	CGA	AGC	TGG	TAC	CTC	ACA	GAG	1824		
Ile	Leu	Phe	Ser	Val	Phe	Asp	Glu	Asn	Arg	Ser	Trp	Tyr	Leu	Thr	Glu			
595			600			605												
AAT	ATA	CAA	CGC	TTT	CTC	CCC	AAT	CCC	GCT	GGA	GTG	CAG	CTT	GAG	GAT	1872		
Asn	Ile	Gln	Arg	Phe	Leu	Pro	Asn	Pro	Ala	Gly	Val	Gln	Leu	Glu	Asp			
610			615			620												
CCA	GAG	TTC	CAA	GCC	TCC	AAC	ATC	ATG	CAC	AGC	ATC	AAT	GGC	TAT	GTT	1920		
Pro	Glu	Phe	Gln	Ala	Ser	Asn	Ile	Met	His	Ser	Ile	Asn	Gly	Tyr	Val			
625			630			635												
TTC	GAT	AGT	TTG	CAG	TTG	TCA	GTT	TGT	TTG	CAT	GAA	GTA	GCA	TAC	TGG	1968		
Phe	Asp	Ser	Leu	Gln	Leu	Ser	Val	Cys	Leu	His	Glu	Val	Ala	Tyr	Trp			
645			650			655												
TAC	ATT	CTA	AGC	ATT	GGA	GCA	CAG	ACT	GAC	TTC	CTT	TCT	GTC	TTC	TTC	2016		
Tyr	Ile	Leu	Ser	Ile	Gly	Ala	Gln	Thr	Asp	Phe	Leu	Ser	Val	Phe	Phe			
660			665			670												

TCT GGA TAT ACC TTC AAA CAC AAA ATG GTC TAT GAA GAC ACA CTC ACC	2064
Ser Gly Tyr Thr Phe Lys His Lys Met Val Tyr Glu Asp Thr Leu Thr	
675 680 685	
CTA TTC CCA TTC TCC GGA GAA ACT GTC TTC ATG TCG ATG GAA AAC CCA	2112
Leu Phe Pro Phe Ser Gly Glu Thr Val Phe Met Ser Met Glu Asn Pro	
690 695 700	
GGA CTA TGG ATT CTG GGG TGC CAC AAC TCA GAC TTT CGG AAC AGA GGC	2160
Gly Leu Trp Ile Leu Gly Cys His Asn Ser Asp Phe Arg Asn Arg Gly	
705 710 715 720	
ATG ACC GCC TTA CTG AAA GTT TCC AGT TGT GAC AAG AAC ACT GGA GAT	2208
Met Thr Ala Leu Leu Lys Val Ser Ser Cys Asp Lys Asn Thr Gly Asp	
725 730 735	
TAT TAC GAG GAC AGT TAT GAA GAT ATT TCA GCA TAC TTG CTG AGT AAA	2256
Tyr Tyr Glu Asp Ser Tyr Glu Asp Ile Ser Ala Tyr Leu Leu Ser Lys	
740 745 750	
AAC AAT GCC ATT GAA CCA AGA AGC TTC TCC CAG AAC CCA CCA GTC TTG	2304
Asn Asn Ala Ile Glu Pro Arg Ser Phe Ser Gln Asn Pro Pro Val Leu	
755 760 765	
AAA CGC CAT CAA CGG GAA ATA ACT CGT ACT ACT CTT CAA TCA GAT CAA	2352
Lys Arg His Gln Arg Glu Ile Thr Arg Thr Thr Leu Gln Ser Asp Gln	
770 775 780	
GAG GAA ATT GAC TAT GAT GAT ACC ATA TCA GTT GAA ATG AAG AAG GAA	2400
Glu Glu Ile Asp Tyr Asp Asp Thr Ile Ser Val Glu Met Lys Lys Glu	
785 790 795 800	
GAT TTC GAC ATT TAT GAT GAG GAT GAA AAT CAG AGC CCC CGC AGC TTT	2448
Asp Phe Asp Ile Tyr Asp Glu Asp Glu Asn Gln Ser Pro Arg Ser Phe	
805 810 815	
CAA AAG AAA ACA CGA CAC TAT TTT ATT GCT GCA GTG GAG AGG CTC TGG	2496
Gln Lys Lys Thr Arg His Tyr Phe Ile Ala Ala Val Glu Arg Leu Trp	
820 825 830	
GAT TAT GGG ATG AGT AGC TCC CCA CAT GTT CTA AGA AAC AGG GCT CAG	2544
Asp Tyr Gly Met Ser Ser Ser Pro His Val Leu Arg Asn Arg Ala Gln	
835 840 845	
AGT GGC AGT GTC CCT CAG TTC AAG AAA GTA GTA TTC CAG GAA TTT ACC	2592
Ser Gly Ser Val Pro Gln Phe Lys Lys Val Val Phe Gln Glu Phe Thr	
850 855 860	
GAT GGC TCC TTT ACT CAA CCC TTA TAC CGT GGA GAA CTA AAT GAA CAT	2640
Asp Gly Ser Phe Thr Gln Pro Leu Tyr Arg Gly Glu Leu Asn Glu His	
865 870 875 880	
TTG GGA CTC CTG GGG CCA TAT ATA AGA GCA GAA GTT GAA GAT AAT ATC	2688
Leu Gly Leu Leu Gly Pro Tyr Ile Arg Ala Glu Val Glu Asp Asn Ile	
885 890 895	

ATG GTT ACC TTC AGA AAT CAG GCC TCT CGT CCC TAT TCC TTC TAT TCT	2736
Met Val Thr Phe Arg Asn Gln Ala Ser Arg Pro Tyr Ser Phe Tyr Ser	
900 905 910	
TCC CTC ATA TCA TAT GAG GAA GAT CAG AGG CAA GGA GCA GAA CCT AGA	2784
Ser Leu Ile Ser Tyr Glu Glu Asp Gln Arg Gln Gly Ala Glu Pro Arg	
915 920 925	
AAA AAC TTT GTC AAG CCT AAT GAA ACC AAA ACT TAC TTT TGG AAA GTG	2832
Lys Asn Phe Val Lys Pro Asn Glu Thr Lys Thr Tyr Phe Trp Lys Val	
930 935 940	
CAA CAT CAT ATG GCA CCC ACT AAA GAT GAG TTT GAC TGC AAA GCC TGG	2880
Gln His His Met Ala Pro Thr Lys Asp Glu Phe Asp Cys Lys Ala Trp	
945 950 955 960	
GCT TAT TTC TCC GAT GTC GAC CTG GAA AAA GAT GTG CAC TCA GGC CTG	2928
Ala Tyr Phe Ser Asp Val Asp Leu Glu Lys Asp Val His Ser Gly Leu	
965 970 975	
ATT GGA CCC CTT CTG GTC TGC CAC ACC AAC ACA CTG AAC CCT GCT CAT	2976
Ile Gly Pro Leu Leu Val Cys His Thr Asn Thr Leu Asn Pro Ala His	
980 985 990	
GGG AGA CAA GTG ACA GTA CAG GAA TTT GCT CTG TTT TTC ACC ATC TTC	3024
Gly Arg Gln Val Thr Val Gln Glu Phe Ala Leu Phe Phe Thr Ile Phe	
995 1000 1005	
GAT GAG ACC AAA AGC TGG TAC TTC ACT GAA AAT ATG GAA AGA AAC TGC	3072
Asp Glu Thr Lys Ser Trp Tyr Phe Thr Glu Asn Met Glu Arg Asn Cys	
1010 1015 1020	
AGG GCT CCC TGC AAT ATC CAG ATG GAA GAT CCC ACT TTT AAA GAG AAT	3120
Arg Ala Pro Cys Asn Ile Gln Met Glu Asp Pro Thr Phe Lys Glu Asn	
1025 1030 1035 1040	
TAT CGC TTC CAT GCA ATC AAT GGC TAC ATA ATG GAT ACA CTA CCT GGC	3168
Tyr Arg Phe His Ala Ile Asn Gly Tyr Ile Met Asp Thr Leu Pro Gly	
1045 1050 1055	
TTA GTA ATG GCT CAG GAT CAA AGG ATT CGA TGG TAT CTG CTC AGC ATG	3216
Leu Val Met Ala Gln Asp Gln Arg Ile Arg Trp Tyr Leu Leu Ser Met	
1060 1065 1070	
GGC AGC AAT GAA AAC ATC CAT TCT ATT CAT TTC TCC GGA CAT GTG TTC	3264
Gly Ser Asn Glu Asn Ile His Ser Ile His Phe Ser Gly His Val Phe	
1075 1080 1085	
ACT GTA CGA AAA AAA GAG GAG TAT AAA ATG GCA CTG TAC AAT CTC TAT	3312
Thr Val Arg Lys Lys Glu Glu Tyr Lys Met Ala Leu Tyr Asn Leu Tyr	
1090 1095 1100	
CCC GGA GTT TTC GAG ACA GTG GAA ATG TTA CCA TCC AAA GCT GGA ATT	3360
Pro Gly Val Phe Glu Thr Val Glu Met Leu Pro Ser Lys Ala Gly Ile	
1105 1110 1115 1120	
TGG CGG GTG GAA TGC CTT ATT GGC GAG CAT CTA CAT GCT GGG ATG AGC	3408

Trp Arg Val Glu Cys Leu Ile Gly Glu His Leu His Ala Gly Met Ser	
1125 1130 1135	
ACA CTT TTT CTG GTG TAC TCC AAT AAG TGT CAG ACT CCC CTG GGA ATG	3456
Thr Leu Phe Leu Val Tyr Ser Asn Lys Cys Gln Thr Pro Leu Gly Met	
1140 1145 1150	
GCT TCT GGA CAC ATT AGA GAT TTT CAG ATT ACA GCT TCA GGA CAA TAT	3504
Ala Ser Gly His Ile Arg Asp Phe Gln Ile Thr Ala Ser Gly Gln Tyr	
1155 1160 1165	
GGA CAG TGG GCC CCA AAG CTG GCC AGA CTT CAT TAT TCC GGA TCA ATC	3552
Gly Gln Trp Ala Pro Lys Leu Ala Arg Leu His Tyr Ser Gly Ser Ile	
1170 1175 1180	
AAT GCC TGG AGC ACC AAG GAG CCC TTT TCT TGG ATC AAA GTT GAC CTG	3600
Asn Ala Trp Ser Thr Lys Glu Pro Phe Ser Trp Ile Lys Val Asp Leu	
1185 1190 1195 1200	
TTG GCA CCA ATG ATT ATT CAC GGC ATC AAG ACC CAG GGT GCC CGT CAG	3648
Leu Ala Pro Met Ile Ile His Gly Ile Lys Thr Gln Gly Ala Arg Gln	
1205 1210 1215	
AAG TTC TCC AGC CTC TAC ATC TCT CAA TTT ATC ATC ATG TAT AGT CTC	3696
Lys Phe Ser Ser Leu Tyr Ile Ser Gln Phe Ile Ile Met Tyr Ser Leu	
1220 1225 1230	
GAT GGG AAG AAG TGG CAG ACT TAT CGA GGA AAT TCC ACT GGA ACC CTC	3744
Asp Gly Lys Lys Trp Gln Thr Tyr Arg Gly Asn Ser Thr Gly Thr Leu	
1235 1240 1245	
ATG GTC TTC TTT GGC AAT GTG GAT TCA TCT GGG ATA AAA CAC AAT ATT	3792
Met Val Phe Phe Gly Asn Val Asp Ser Ser Gly Ile Lys His Asn Ile	
1250 1255 1260	
TTC AAC CCT CCA ATT ATT GCT CGA TAC ATC CGT TTG CAC CCA ACT CAT	3840
Phe Asn Pro Pro Ile Ile Ala Arg Tyr Ile Arg Leu His Pro Thr His	
1265 1270 1275 1280	
TAT AGC ATT CGC AGC ACT CTT CGC ATG GAG TTG ATG GGC TGT GAT TTA	3888
Tyr Ser Ile Arg Ser Thr Leu Arg Met Glu Leu Met Gly Cys Asp Leu	
1285 1290 1295	
AAT AGT TGC AGC ATG CCA TTG GGA ATG GAG AGT AAA GCA ATA TCA GAT	3936
Asn Ser Cys Ser Met Pro Leu Gly Met Glu Ser Lys Ala Ile Ser Asp	
1300 1305 1310	
GCA CAG ATT ACT GCT TCA TCC TAC TTT ACC AAT ATG TTT GCC ACC TGG	3984
Ala Gln Ile Thr Ala Ser Ser Tyr Phe Thr Asn Met Phe Ala Thr Trp	
1315 1320 1325	
TCT CCT TCA AAA GCT CGA CTA CAC CTA CAA GGG AGG AGT AAT GCC TGG	4032
Ser Pro Ser Lys Ala Arg Leu His Leu Gln Gly Arg Ser Asn Ala Trp	
1330 1335 1340	
AGA CCT CAA GTT AAC AAT CCA AAA GAG TGG CTG CAA GTG GAC TTC CAG	4080
Arg Pro Gln Val Asn Asn Pro Lys Glu Trp Leu Gln Val Asp Phe Gln	

1345	1350	1355	1360	
AAG ACA ATG AAA GTC ACA GGA GTA ACT ACT CAG GGA GTA AAA TCT CTG				4128
Lys Thr Met Lys Val Thr Gly Val Thr Thr Gln Gly Val Lys Ser Leu				
	1365	1370	1375	
CTT ACC TCT ATG TAC GTG AAG GAG TTC CTC ATA TCG TCG TCG CAA GAT				4176
Leu Thr Ser Met Tyr Val Lys Glu Phe Leu Ile Ser Ser Ser Gln Asp				
	1380	1385	1390	
GGC CAT CAG TGG ACT CTC TTT TTT CAA AAT GGC AAA GTA AAA GTT TTC				4224
Gly His Gln Trp Thr Leu Phe Phe Gln Asn Gly Lys Val Lys Val Phe				
	1395	1400	1405	
CAG GGA AAT CAA GAC TCC TTC ACA CCT GTC GTG AAC TCT CTA GAC CCA				4272
Gln Gly Asn Gln Asp Ser Phe Thr Pro Val Val Asn Ser Leu Asp Pro				
	1410	1415	1420	
CCG TTA CTC ACT CGC TAC CTT CGA ATT CAC CCC CAG AGT TGG GTG CAC				4320
Pro Leu Leu Thr Arg Tyr Leu Arg Ile His Pro Gln Ser Trp Val His				
	1425	1430	1435	1440
CAG ATT GCC CTG AGG ATG GAG GTT CTG GGC TGC GAG GCA CAG GAC CTC				4368
Gln Ile Ala Leu Arg Met Glu Val Leu Gly Cys Glu Ala Gln Asp Leu				
	1445	1450	1455	
TAC TG				4374
Tyr				

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 9164 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1006..5376

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

GTCGACGGTA TCGATAAGCT TGATATCGAA TTCCTGCAGC CCGGGGGATC CACTAGTACT	60
CGAGACCTAG GAGTTAATTT TTAAAAAGCA GTCAAAAGTC CAAGTGGCCC TTGCGAGCAT	120
TTACTCTCTC TGTTTGCTCT GGTAAATAAT CTCAGGAGCA CAAACATTCC TTACTAGTCC	180
TAGAAGTTAA TTTTAAAAA GCAGTCAAAA GTCCAAGTGG CCCTTGCGAG CATTTACTCT	240
CTCTGTTTGC TCTGGTTAAT AATCTCAGGA GCACAAACAT TCCTTACTAG TTCTAGAGCG	300
GCCGCCAGTG TGCTGGAATT CGGCTTTTTT AGGGCTGGAA GCTACCTTTG ACATCATTTTC	360

CTCTGCGAAT GCATGTATAA TTTCTACAGA ACCTATTAGA AAGGATCACC CAGCCTCTGC	420
TTTTGTACAA CTTTCCCTTA AAAAAGTACC AATCCACTG CTGTTTGGCC CAATAGTGAG	480
AACTTTTTCC TGCTGCCTCT TGGTGCTTTT GCCTATGGCC CCTATTCTGC CTGCTGAAGA	540
CACTCTTGCC AGCATGGACT TAAACCCCTC CAGCTCTGAC AATCCTCTTT CTCTTTTGTT	600
TTACATGAAG GGTCTGGCAG CCAAAGCAAT CACTCAAAGT TCAAACCTTA TCATTTTTTG	660
CTTTGTTCCT CTTGGCCTTG GTTTTGTACA TCAGCTTTGA AAATACCATC CCAGGGTTAA	720
TGCTGGGGTT AATTTATAAC TAAGAGTGCT CTAGTTTGC AATACAGGAC ATGCTATAAA	780
AATGAAAAGA TGTGCTTTC TGAGAGATCT CGAGGAAGCT AACAACAAAG AACAACAAAC	840
AACAATCAGG TAAGTATCCT TTTTACAGCA CAACTTAATG AGACAGATAG AAAGTGGTCT	900
TGTAGAAACA GAGTAGTCGC CTGCTTTTCT GCCAGGTGCT GACTTCTCTC CCCTTCTCTT	960
TTTTCTTTT CTCAGGATAA CAAGAACGAA ACAATAACAG CCACC ATG GAA ATA	1014
Met Glu Ile	
1	
GAG CTC TCC ACC TGC TTC TTT CTG TGC CTT TTG CGA TTC TGC TTT AGT	1062
Glu Leu Ser Thr Cys Phe Phe Leu Cys Leu Leu Arg Phe Cys Phe Ser	
5 10 15	
GCC ACC AGA AGA TAC TAC CTG GGT GCA GTG GAA CTG TCA TGG GAC TAT	1110
Ala Thr Arg Arg Tyr Tyr Leu Gly Ala Val Glu Leu Ser Trp Asp Tyr	
20 25 30 35	
ATG CAA AGT GAT CTC GGT GAG CTG CCT GTG GAC GCA AGA TTT CCT CCT	1158
Met Gln Ser Asp Leu Gly Glu Leu Pro Val Asp Ala Arg Phe Pro Pro	
40 45 50	
AGA GTG CCA AAA TCT TTT CCA TTC AAC ACC TCA GTC GTG TAC AAA AAG	1206
Arg Val Pro Lys Ser Phe Pro Phe Asn Thr Ser Val Val Tyr Lys Lys	
55 60 65	
ACT CTG TTT GTA GAA TTC ACG GTT CAC CTT TTC AAC ATC GCT AAG CCA	1254
Thr Leu Phe Val Glu Phe Thr Val His Leu Phe Asn Ile Ala Lys Pro	
70 75 80	
AGG CCA CCC TGG ATG GGT CTG CTA GGT CCT ACC ATC CAG GCT GAG GTT	1302
Arg Pro Pro Trp Met Gly Leu Leu Gly Pro Thr Ile Gln Ala Glu Val	
85 90 95	
TAT GAT ACA GTG GTC ATT ACA CTT AAG AAC ATG GCT TCC CAT CCT GTC	1350
Tyr Asp Thr Val Val Ile Thr Leu Lys Asn Met Ala Ser His Pro Val	
100 105 110 115	
AGT CTT CAT GCT GTT GGT GTA TCC TAC TGG AAA GCT TCT GAG GGA GCT	1398
Ser Leu His Ala Val Gly Val Ser Tyr Trp Lys Ala Ser Glu Gly Ala	
120 125 130	
GAA TAT GAT GAT CAG ACC AGT CAA AGG GAG AAA GAA GAT GAT AAA GTC	1446

Glu Tyr Asp Asp Gln Thr Ser Gln Arg Glu Lys Glu Asp Asp Lys Val	
135	140
145	
TTC CCT GGT GGA AGC CAT ACA TAT GTC TGG CAG GTC CTG AAA GAG AAT	1494
Phe Pro Gly Gly Ser His Thr Tyr Val Trp Gln Val Leu Lys Glu Asn	
150	155
160	
GGT CCA ATG GCC TCT GAC CCA CTG TGC CTT ACC TAC TCA TAT CTT TCT	1542
Gly Pro Met Ala Ser Asp Pro Leu Cys Leu Thr Tyr Ser Tyr Leu Ser	
165	170
175	
CAT GTG GAC CTG GTA AAA GAC TTG AAT TCA GGC CTC ATT GGA GCC CTA	1590
His Val Asp Leu Val Lys Asp Leu Asn Ser Gly Leu Ile Gly Ala Leu	
180	185
190	195
CTA GTA TGT AGA GAA GGG AGT CTG GCC AAG GAA AAG ACA CAG ACC TTG	1638
Leu Val Cys Arg Glu Gly Ser Leu Ala Lys Glu Lys Thr Gln Thr Leu	
200	205
210	
CAC AAA TTT ATA CTA CTT TTT GCT GTA TTT GAT GAA GGG AAA AGT TGG	1686
His Lys Phe Ile Leu Leu Phe Ala Val Phe Asp Glu Gly Lys Ser Trp	
215	220
225	
CAC TCA GAA ACA AAG AAC TCC TTG ATG CAG GAT AGG GAT GCT GCA TCT	1734
His Ser Glu Thr Lys Asn Ser Leu Met Gln Asp Arg Asp Ala Ala Ser	
230	235
240	
GCT CGG GCC TGG CCT AAA ATG CAC ACA GTC AAT GGT TAT GTA AAC AGG	1782
Ala Arg Ala Trp Pro Lys Met His Thr Val Asn Gly Tyr Val Asn Arg	
245	250
255	
TCT CTG CCA GGT CTG ATT GGA TGC CAC AGG AAA TCA GTC TAT TGG CAT	1830
Ser Leu Pro Gly Leu Ile Gly Cys His Arg Lys Ser Val Tyr Trp His	
260	265
270	275
GTG ATT GGA ATG GGC ACC ACT CCT GAA GTG CAC TCA ATA TTC CTC GAA	1878
Val Ile Gly Met Gly Thr Thr Pro Glu Val His Ser Ile Phe Leu Glu	
280	285
290	
GGT CAC ACA TTT CTT GTG AGG AAC CAT CGC CAG GCG TCC TTG GAA ATC	1926
Gly His Thr Phe Leu Val Arg Asn His Arg Gln Ala Ser Leu Glu Ile	
295	300
305	
TCG CCA ATA ACT TTC CTT ACT GCT CAA ACA CTC TTG ATG GAC CTT GGA	1974
Ser Pro Ile Thr Phe Leu Thr Ala Gln Thr Leu Leu Met Asp Leu Gly	
310	315
320	
CAG TTT CTA CTG TTT TGT CAT ATC TCT TCC CAC CAA CAT GAT GGC ATG	2022
Gln Phe Leu Leu Phe Cys His Ile Ser Ser His Gln His Asp Gly Met	
325	330
335	
GAA GCT TAT GTC AAA GTA GAC AGC TGT CCA GAG GAA CCC CAA CTA CGA	2070
Glu Ala Tyr Val Lys Val Asp Ser Cys Pro Glu Glu Pro Gln Leu Arg	
340	345
350	355
ATG AAA AAT AAT GAA GAA GCG GAA GAC TAT GAT GAT GAT CTT ACT GAT	2118
Met Lys Asn Asn Glu Glu Ala Glu Asp Tyr Asp Asp Asp Leu Thr Asp	

360										365					370					
TCT	GAA	ATG	GAT	GTG	GTC	AGG	TTT	GAT	GAT	GAC	AAC	TCT	CCT	TCC	TTT		2166			
Ser	Glu	Met	Asp	Val	Val	Arg	Phe	Asp	Asp	Asp	Asn	Ser	Pro	Ser	Phe					
375						380					385									
ATC	CAA	ATT	CGC	TCA	GTT	GCC	AAG	AAG	CAT	CCT	AAA	ACT	TGG	GTA	CAT		2214			
Ile	Gln	Ile	Arg	Ser	Val	Ala	Lys	Lys	His	Pro	Lys	Thr	Trp	Val	His					
390						395					400									
TAC	ATT	GCT	GCT	GAA	GAG	GAG	GAC	TGG	GAC	TAT	GCT	CCC	TTA	GTC	CTC		2262			
Tyr	Ile	Ala	Ala	Glu	Glu	Glu	Asp	Trp	Asp	Tyr	Ala	Pro	Leu	Val	Leu					
405						410					415									
GCC	CCC	GAT	GAC	AGA	AGT	TAT	AAA	AGT	CAA	TAT	TTG	AAC	AAT	GGC	CCT		2310			
Ala	Pro	Asp	Asp	Arg	Ser	Tyr	Lys	Ser	Gln	Tyr	Leu	Asn	Asn	Gly	Pro					
420						425					430					435				
CAG	CGG	ATT	GGT	AGG	AAG	TAC	AAA	AAA	GTC	CGA	TTT	ATG	GCA	TAC	ACA		2358			
Gln	Arg	Ile	Gly	Arg	Lys	Tyr	Lys	Lys	Val	Arg	Phe	Met	Ala	Tyr	Thr					
440						445					450									
GAT	GAA	ACC	TTT	AAG	ACT	CGT	GAA	GCT	ATT	CAG	CAT	GAA	TCA	GGA	ATC		2406			
Asp	Glu	Thr	Phe	Lys	Thr	Arg	Glu	Ala	Ile	Gln	His	Glu	Ser	Gly	Ile					
455						460					465									
TTG	GGA	CCT	TTA	CTT	TAT	GGG	GAA	GTT	GGA	GAC	ACA	CTG	TTG	ATT	ATA		2454			
Leu	Gly	Pro	Leu	Leu	Tyr	Gly	Glu	Val	Gly	Asp	Thr	Leu	Leu	Ile	Ile					
470						475					480									
TTT	AAG	AAT	CAA	GCA	AGC	AGA	CCA	TAT	AAC	ATC	TAC	CCT	CAC	GGA	ATC		2502			
Phe	Lys	Asn	Gln	Ala	Ser	Arg	Pro	Tyr	Asn	Ile	Tyr	Pro	His	Gly	Ile					
485						490					495									
ACT	GAT	GTC	CGT	CCT	TTG	TAT	TCA	AGG	AGA	TTA	CCA	AAA	GGT	GTA	AAA		2550			
Thr	Asp	Val	Arg	Pro	Leu	Tyr	Ser	Arg	Arg	Leu	Pro	Lys	Gly	Val	Lys					
500						505					510					515				
CAT	TTG	AAG	GAT	TTT	CCA	ATT	CTG	CCA	GGA	GAA	ATA	TTC	AAA	TAT	AAA		2598			
His	Leu	Lys	Asp	Phe	Pro	Ile	Leu	Pro	Gly	Glu	Ile	Phe	Lys	Tyr	Lys					
520						525					530									
TGG	ACA	GTG	ACT	GTA	GAA	GAT	GGG	CCA	ACT	AAA	TCA	GAT	CCT	CGG	TGC		2646			
Trp	Thr	Val	Thr	Val	Glu	Asp	Gly	Pro	Thr	Lys	Ser	Asp	Pro	Arg	Cys					
535						540					545									
CTG	ACC	CGC	TAT	TAC	TCT	AGT	TTC	GTT	AAT	ATG	GAG	AGA	GAT	CTA	GCT		2694			
Leu	Thr	Arg	Tyr	Tyr	Ser	Ser	Phe	Val	Asn	Met	Glu	Arg	Asp	Leu	Ala					
550						555					560									
TCA	GGA	CTC	ATT	GGC	CCT	CTC	CTC	ATC	TGC	TAC	AAA	GAA	TCT	GTA	GAT		2742			
Ser	Gly	Leu	Ile	Gly	Pro	Leu	Leu	Ile	Cys	Tyr	Lys	Glu	Ser	Val	Asp					
565						570					575									
CAA	AGA	GGA	AAC	CAG	ATA	ATG	TCA	GAC	AAG	AGG	AAT	GTC	ATC	CTG	TTT		2790			
Gln	Arg	Gly	Asn	Gln	Ile	Met	Ser	Asp	Lys	Arg	Asn	Val	Ile	Leu	Phe					
580						585					590					595				

TCT GTA TTT GAT GAG AAC CGA AGC TGG TAC CTC ACA GAG AAT ATA CAA	2838
Ser Val Phe Asp Glu Asn Arg Ser Trp Tyr Leu Thr Glu Asn Ile Gln	
600 605 610	
CGC TTT CTC CCC AAT CCA GCT GGA GTG CAG CTT GAG GAT CCA GAG TTC	2886
Arg Phe Leu Pro Asn Pro Ala Gly Val Gln Leu Glu Asp Pro Glu Phe	
615 620 625	
CAA GCC TCC AAC ATC ATG CAC AGC ATC AAT GGC TAT GTT TTT GAT AGT	2934
Gln Ala Ser Asn Ile Met His Ser Ile Asn Gly Tyr Val Phe Asp Ser	
630 635 640	
TTG CAG TTG TCA GTT TGT TTG CAT GAG GTG GCA TAC TGG TAC ATT CTA	2982
Leu Gln Leu Ser Val Cys Leu His Glu Val Ala Tyr Trp Tyr Ile Leu	
645 650 655	
AGC ATT GGA GCA CAG ACT GAC TTC CTT TCT GTC TTC TTC TCT GGA TAT	3030
Ser Ile Gly Ala Gln Thr Asp Phe Leu Ser Val Phe Phe Ser Gly Tyr	
660 665 670 675	
ACC TTC AAA CAC AAA ATG GTC TAT GAA GAC ACA CTC ACC CTA TTC CCA	3078
Thr Phe Lys His Lys Met Val Tyr Glu Asp Thr Leu Thr Leu Phe Pro	
680 685 690	
TTC TCA GGA GAA ACT GTC TTC ATG TCG ATG GAA AAC CCA GGT CTA TGG	3126
Phe Ser Gly Glu Thr Val Phe Met Ser Met Glu Asn Pro Gly Leu Trp	
695 700 705	
ATT CTG GGG TGC CAC AAC TCA GAC TTT CGG AAC AGA GGC ATG ACC GCC	3174
Ile Leu Gly Cys His Asn Ser Asp Phe Arg Asn Arg Gly Met Thr Ala	
710 715 720	
TTA CTG AAG GTT TCT AGT TGT GAC AAG AAC ACT GGT GAT TAT TAC GAG	3222
Leu Leu Lys Val Ser Ser Cys Asp Lys Asn Thr Gly Asp Tyr Tyr Glu	
725 730 735	
GAC AGT TAT GAA GAT ATT TCA GCA TAC TTG CTG AGT AAA AAC AAT GCC	3270
Asp Ser Tyr Glu Asp Ile Ser Ala Tyr Leu Leu Ser Lys Asn Asn Ala	
740 745 750 755	
ATT GAA CCA AGA AGC TTC TCC CAG AAC CCA CCA GTC TTG AAA CGC CAT	3318
Ile Glu Pro Arg Ser Phe Ser Gln Asn Pro Pro Val Leu Lys Arg His	
760 765 770	
CAA CGG GAA ATA ACT CGT ACT ACT CTT CAG TCA GAT CAA GAG GAA ATT	3366
Gln Arg Glu Ile Thr Arg Thr Thr Leu Gln Ser Asp Gln Glu Glu Ile	
775 780 785	
GAC TAT GAT GAT ACC ATA TCA GTT GAA ATG AAG AAG GAA GAT TTT GAC	3414
Asp Tyr Asp Asp Thr Ile Ser Val Glu Met Lys Lys Glu Asp Phe Asp	
790 795 800	
ATT TAT GAT GAG GAT GAA AAT CAG AGC CCC CGC AGC TTT CAA AAG AAA	3462
Ile Tyr Asp Glu Asp Glu Asn Gln Ser Pro Arg Ser Phe Gln Lys Lys	
805 810 815	

ACA CGA CAC TAT TTT ATT GCT GCA GTG GAG AGG CTC TGG GAT TAT GGG	3510
Thr Arg His Tyr Phe Ile Ala Ala Val Glu Arg Leu Trp Asp Tyr Gly	
820 825 830 835	
ATG AGT AGC TCC CCA CAT GTT CTA AGA AAC AGG GCT CAG AGT GGC AGT	3558
Met Ser Ser Ser Pro His Val Leu Arg Asn Arg Ala Gln Ser Gly Ser	
840 845 850	
GTC CCT CAG TTC AAG AAA GTT GTT TTC CAG GAA TTT ACT GAT GGC TCC	3606
Val Pro Gln Phe Lys Lys Val Val Phe Gln Glu Phe Thr Asp Gly Ser	
855 860 865	
TTT ACT CAG CCC TTA TAC CGT GGA GAA CTA AAT GAA CAT TTG GGA CTC	3654
Phe Thr Gln Pro Leu Tyr Arg Gly Glu Leu Asn Glu His Leu Gly Leu	
870 875 880	
CTG GGG CCA TAT ATA AGA GCA GAA GTT GAA GAT AAT ATC ATG GTA ACT	3702
Leu Gly Pro Tyr Ile Arg Ala Glu Val Glu Asp Asn Ile Met Val Thr	
885 890 895	
TTC AGA AAT CAG GCC TCT CGT CCC TAT TCC TTC TAT TCT AGC CTT ATT	3750
Phe Arg Asn Gln Ala Ser Arg Pro Tyr Ser Phe Tyr Ser Ser Leu Ile	
900 905 910 915	
TCT TAT GAG GAA GAT CAG AGG CAA GGA GCA GAA CCT AGA AAA AAC TTT	3798
Ser Tyr Glu Glu Asp Gln Arg Gln Gly Ala Glu Pro Arg Lys Asn Phe	
920 925 930	
GTC AAG CCT AAT GAA ACC AAA ACT TAC TTT TGG AAA GTG CAA CAT CAT	3846
Val Lys Pro Asn Glu Thr Lys Thr Tyr Phe Trp Lys Val Gln His His	
935 940 945	
ATG GCA CCC ACT AAA GAT GAG TTT GAC TGC AAA GCC TGG GCT TAT TTC	3894
Met Ala Pro Thr Lys Asp Glu Phe Asp Cys Lys Ala Trp Ala Tyr Phe	
950 955 960	
TCT GAT GTT GAC CTG GAA AAA GAT GTG CAC TCA GGC CTG ATT GGA CCC	3942
Ser Asp Val Asp Leu Glu Lys Asp Val His Ser Gly Leu Ile Gly Pro	
965 970 975	
CTT CTG GTC TGC CAC ACT AAC ACA CTG AAC CCT GCT CAT GGG AGA CAA	3990
Leu Leu Val Cys His Thr Asn Thr Leu Asn Pro Ala His Gly Arg Gln	
980 985 990 995	
GTG ACA GTA CAG GAA TTT GCT CTG TTT TTC ACC ATC TTT GAT GAG ACC	4038
Val Thr Val Gln Glu Phe Ala Leu Phe Phe Thr Ile Phe Asp Glu Thr	
1000 1005 1010	
AAA AGC TGG TAC TTC ACT GAA AAT ATG GAA AGA AAC TGC AGG GCT CCC	4086
Lys Ser Trp Tyr Phe Thr Glu Asn Met Glu Arg Asn Cys Arg Ala Pro	
1015 1020 1025	
TGC AAT ATC CAG ATG GAA GAT CCC ACT TTT AAA GAG AAT TAT CGC TTC	4134
Cys Asn Ile Gln Met Glu Asp Pro Thr Phe Lys Glu Asn Tyr Arg Phe	
1030 1035 1040	
CAT GCA ATC AAT GGC TAC ATA ATG GAT ACA CTA CCT GGC TTA GTA ATG	4182

His Ala Ile Asn Gly Tyr Ile Met Asp Thr Leu Pro Gly Leu Val Met	
1045 1050 1055	
GCT CAG GAT CAA AGG ATT CGA TGG TAT CTG CTC AGC ATG GGC AGC AAT	4230
Ala Gln Asp Gln Arg Ile Arg Trp Tyr Leu Leu Ser Met Gly Ser Asn	
1060 1065 1070 1075	
GAA AAC ATC CAT TCT ATT CAT TTC AGT GGA CAT GTG TTC ACT GTA CGA	4278
Glu Asn Ile His Ser Ile His Phe Ser Gly His Val Phe Thr Val Arg	
1080 1085 1090	
AAA AAA GAG GAG TAT AAA ATG GCA CTG TAC AAT CTC TAT CCA GGT GTT	4326
Lys Lys Glu Glu Tyr Lys Met Ala Leu Tyr Asn Leu Tyr Pro Gly Val	
1095 1100 1105	
TTT GAG ACA GTG GAA ATG TTA CCA TCC AAA GCT GGA ATT TGG CGG GTG	4374
Phe Glu Thr Val Glu Met Leu Pro Ser Lys Ala Gly Ile Trp Arg Val	
1110 1115 1120	
GAA TGC CTT ATT GGC GAG CAT CTA CAT GCT GGG ATG AGC ACA CTT TTT	4422
Glu Cys Leu Ile Gly Glu His Leu His Ala Gly Met Ser Thr Leu Phe	
1125 1130 1135	
CTG GTG TAC AGC AAT AAG TGT CAG ACT CCC CTG GGA ATG GCT TCT GGA	4470
Leu Val Tyr Ser Asn Lys Cys Gln Thr Pro Leu Gly Met Ala Ser Gly	
1140 1145 1150 1155	
CAC ATT AGA GAT TTT CAG ATT ACA GCT TCA GGA CAA TAT GGA CAG TGG	4518
His Ile Arg Asp Phe Gln Ile Thr Ala Ser Gly Gln Tyr Gly Gln Trp	
1160 1165 1170	
GCC CCA AAG CTG GCC AGA CTT CAT TAT TCC GGA TCA ATC AAT GCC TGG	4566
Ala Pro Lys Leu Ala Arg Leu His Tyr Ser Gly Ser Ile Asn Ala Trp	
1175 1180 1185	
AGC ACC AAG GAG CCC TTT TCT TGG ATC AAG GTG GAT CTG TTG GCA CCA	4614
Ser Thr Lys Glu Pro Phe Ser Trp Ile Lys Val Asp Leu Leu Ala Pro	
1190 1195 1200	
ATG ATT ATT CAC GGC ATC AAG ACC CAG GGT GCC CGT CAG AAG TTC TCC	4662
Met Ile Ile His Gly Ile Lys Thr Gln Gly Ala Arg Gln Lys Phe Ser	
1205 1210 1215	
AGC CTC TAC ATC TCT CAG TTT ATC ATC ATG TAT AGT CTT GAT GGG AAG	4710
Ser Leu Tyr Ile Ser Gln Phe Ile Ile Met Tyr Ser Leu Asp Gly Lys	
1220 1225 1230 1235	
AAG TGG CAG ACT TAT CGA GGA AAT TCC ACT GGA ACC TTA ATG GTC TTC	4758
Lys Trp Gln Thr Tyr Arg Gly Asn Ser Thr Gly Thr Leu Met Val Phe	
1240 1245 1250	
TTT GGC AAT GTG GAT TCA TCT GGG ATA AAA CAC AAT ATT TTT AAC CCT	4806
Phe Gly Asn Val Asp Ser Ser Gly Ile Lys His Asn Ile Phe Asn Pro	
1255 1260 1265	
CCA ATT ATT GCT CGA TAC ATC CGT TTG CAC CCA ACT CAT TAT AGC ATT	4854
Pro Ile Ile Ala Arg Tyr Ile Arg Leu His Pro Thr His Tyr Ser Ile	

1270	1275	1280	
CGC AGC ACT CTT CGC ATG GAG TTG ATG GGC TGT GAT TTA AAT AGT TGC			4902
Arg Ser Thr Leu Arg Met Glu Leu Met Gly Cys Asp Leu Asn Ser Cys			
1285	1290	1295	
AGC ATG CCA TTG GGA ATG GAG AGT AAA GCA ATA TCA GAT GCA CAG ATT			4950
Ser Met Pro Leu Gly Met Glu Ser Lys Ala Ile Ser Asp Ala Gln Ile			
1300	1305	1310	1315
ACT GCT TCA TCC TAC TTT ACC AAT ATG TTT GCC ACC TGG TCT CCT TCA			4998
Thr Ala Ser Ser Tyr Phe Thr Asn Met Phe Ala Thr Trp Ser Pro Ser			
1320	1325	1330	
AAA GCT CGA CTT CAC CTC CAA GGG AGG AGT AAT GCC TGG AGA CCT CAG			5046
Lys Ala Arg Leu His Leu Gln Gly Arg Ser Asn Ala Trp Arg Pro Gln			
1335	1340	1345	
GTG AAT AAT CCA AAA GAG TGG CTG CAA GTG GAC TTC CAG AAG ACA ATG			5094
Val Asn Asn Pro Lys Glu Trp Leu Gln Val Asp Phe Gln Lys Thr Met			
1350	1355	1360	
AAA GTC ACA GGA GTA ACT ACT CAG GGA GTA AAA TCT CTG CTT ACC AGC			5142
Lys Val Thr Gly Val Thr Thr Gln Gly Val Lys Ser Leu Leu Thr Ser			
1365	1370	1375	
ATG TAT GTG AAG GAG TTC CTC ATC TCC AGC AGT CAA GAT GGC CAT CAG			5190
Met Tyr Val Lys Glu Phe Leu Ile Ser Ser Ser Gln Asp Gly His Gln			
1380	1385	1390	1395
TGG ACT CTC TTT TTT CAG AAT GGC AAA GTA AAG GTT TTT CAG GGA AAT			5238
Trp Thr Leu Phe Phe Gln Asn Gly Lys Val Lys Val Phe Gln Gly Asn			
1400	1405	1410	
CAA GAC TCC TTC ACA CCT GTG GTG AAC TCT CTA GAC CCA CCG TTA CTG			5286
Gln Asp Ser Phe Thr Pro Val Val Asn Ser Leu Asp Pro Pro Leu Leu			
1415	1420	1425	
ACT CGC TAC CTT CGA ATT CAC CCC CAG AGT TGG GTG CAC CAG ATT GCC			5334
Thr Arg Tyr Leu Arg Ile His Pro Gln Ser Trp Val His Gln Ile Ala			
1430	1435	1440	
CTG AGG ATG GAG GTT CTG GGC TGC GAG GCA CAG GAC CTC TAC			5376
Leu Arg Met Glu Val Leu Gly Cys Glu Ala Gln Asp Leu Tyr			
1445	1450	1455	
TGAGGGTGGC CACTGCAGCA CCTGCCACTG CCGTCACCTC TCCCTCCTCA GCTCCAGGGC			5436
AGTGTCCCTC CCTGGCTTGC CTTCTACCTT TGTGCTAAAT CCTAGCAGAC ACTGCCTTGA			5496
AGCCTCCTGA ATTAACATATC ATCAGTCCTG CATTTCTTTG GTGGGGGGGCC AGGAGGGTGC			5556
ATCCAATTTA ACTTAACTCT TACCTATTTT CTGCAGCTGC TCCCAGATTA CTCCTTCCTT			5616
CCAATATAAC TAGGCAAAAA GAAGTGAGGA GAAACCTGCA TGAAAGCATT CTTCCCTGAA			5676
AAGTTAGGCC TCTCAGAGTC ACCACTTCCT CTGTTGTAGA AAAACTATGT GATGAAACTT			5736

TGAAAAAGAT	ATTTATGATG	TTAACTTGTT	TATTGCAGCT	TATAATGGTT	ACAAATAAAG	5796
CAATAGCATC	ACAAATTTCA	CAAATAAAGC	ATTTTTTTTCA	CTGCATTCTA	GTTGTGGTTT	5856
GTCCAAACTC	ATCAATGTAT	CTTATCATGT	CTGGATCCCC	GGGTGGCATC	CCTGTGACCC	5916
CTCCCCAGTG	CCTCTCCTGG	CCCTGGAAGT	TGCCACTCCA	GTGCCCACCA	GCCTTGTCTT	5976
AATAAAATTA	AGTTGCATCA	TTTTGTCTGA	CTAGGTGTCC	TTCTATAATA	TTATGGGGTG	6036
GAGGGGGGTG	GTATGGAGCA	AGGGGCAAGT	TGGGAAGACA	ACCTGTAGGG	CCTGCGGGGT	6096
CTATTGCGGA	ACCAAGCTGG	AGTGCAGTGG	CACAATCTTG	GCTCACTGCA	ATCTCCGCCT	6156
CCTGGGTTC	AGCGATTCTC	CTGCCTCAGC	CTCCCGAGTT	GTTGGGATTC	CAGGCATGCA	6216
TGACCAGGCT	CAGCTAATTT	TTGTTTTTTT	GGTAGAGACG	GGGTTCACC	ATATTGGCCA	6276
GGCTGGTCTC	CAACTCCTAA	TCTCAGGTGA	TCTACCCACC	TTGGCCTCCC	AAATTGCTGG	6336
GATTACAGGC	GTGAACCACT	GCTCCCTTCC	CTGTCTTTCT	GATTTTAAAA	TAATAATACC	6396
AGCAGGAGGA	CGTCCAGACA	CAGCATAGGC	TACCTGCCAT	GCCCAACCGG	TGGGACATTT	6456
GAGTTGCTTG	CTTGGCACTG	TCCTCTCATG	CGTTGGGTCC	ACTCAGTAGA	TGCCTGTTGA	6516
ATTGTAATC	ATGGTCATAG	CTGTTTCCTG	TGTGAAATTG	TTATCCGCTC	ACAATTCCAC	6576
ACAACATACG	AGCCGGAAGC	ATAAAGTGTA	AAGCCTGGGG	TGCCTAATGA	GTGAGCTAAC	6636
TCACATTAAT	TGCGTTGCGC	TCACTGCCCC	CTTTCCAGTC	GGGAAACCTG	TCGTGCCAGC	6696
TGCATTAATG	AATCGGCCAA	CGCGCGGGGA	GAGGCGGTTT	GCGTATTGGG	CGCTCTTCCG	6756
CTTCCTCGCT	CACTGACTCG	CTGCGCTCGG	TCGTTGCGCT	GCGGCGAGCG	GTATCAGCTC	6816
ACTCAAAGGC	GGTAATACGG	TTATCCACAG	AATCAGGGGA	TAACGCAGGA	AAGAACATGT	6876
GAGCAAAAGG	CCAGCAAAAG	GCCAGGAACC	GTAAAAAGGC	CGCGTTGCTG	GCGTTTTTCC	6936
ATAGGCTCCG	CCCCCCTGAC	GAGCATCACA	AAAATCGACG	CTCAAGTCAG	AGGTGGCGAA	6996
ACCCGACAGG	ACTATAAAGA	TACCAGGCGT	TTCCCCCTGG	AAGCTCCCTC	GTGCGCTCTC	7056
CTGTTCCGAC	CCTGCCGCTT	ACCGGATACC	TGTCCGCCTT	TCTCCCTTCG	GGAAGCGTGG	7116
CGCTTTCTCA	TAGCTCACGC	TGTAGGTATC	TCAGTTCGGT	GTAGGTCGTT	CGCTCCAAGC	7176
TGGGCTGTGT	GCACGAACCC	CCCGTTCAGC	CCGACCGCTG	CGCCTTATCC	GGTAACTATC	7236
GTCTTGAGTC	CAACCCGGTA	AGACACGACT	TATCGCCACT	GGCAGCAGCC	ACTGGTAACA	7296
GGATTAGCAG	AGCGAGGTAT	GTAGGCGGTG	CTACAGAGTT	CTTGAAGTGG	TGGCCTAACT	7356
ACGGCTACAC	TAGAAGGACA	GTATTTGGTA	TCTGCGCTCT	GCTGAAGCCA	GTTACCTTCG	7416

GAAAAAGAGT	TGGTAGCTCT	TGATCCGGCA	AACAAACCAC	CGCTGGTAGC	GGTGGTTTTT	7476
TTGTTTGCAA	GCAGCAGATT	ACGCGCAGAA	AAAAAGGATC	TCAAGAAGAT	CCTTTGATCT	7536
TTTCTACGGG	GTCTGACGCT	CAGTGGAACG	AAAAC TCACG	TTAAGGGATT	TTGGTCATGA	7596
GATTATCAAA	AAGGATCTTC	ACCTAGATCC	TTTTAAATTA	AAAATGAAGT	TTTAAATCAA	7656
TCTAAAGTAT	ATATGAGTAA	ACTTGGTCTG	ACAGTTACCA	ATGCTTAATC	AGTGAGGCAC	7716
CTATCTCAGC	GATCTGTCTA	TTTCGTTTAT	CCATAGTTGC	CTGACTCCCC	GTCGTGTAGA	7776
TAACTACGAT	ACGGGAGGGC	TTACCATCTG	GCCCCAGTGC	TGCAATGATA	CCGCGAGACC	7836
CACGCTCACC	GGCTCCAGAT	TTATCAGCAA	TAAACCAGCC	AGCCGGAAGG	GCCGAGCGCA	7896
GAAGTGGTCC	TGCAACTTTA	TCCGCCTCCA	TCCAGTCTAT	TAATTGTTGC	CGGGAAGCTA	7956
GAGTAAGTAG	TTCGCCAGTT	AATAGTTTGC	GCAACGTTGT	TGCCATTGCT	ACAGGCATCG	8016
TGGTGTCACG	CTCGTCGTTT	GGTATGGCTT	CATTCAGCTC	CGGTTCCCAA	CGATCAAGGC	8076
GAGTTACATG	ATCCCCCATG	TTGTGCAAAA	AAGCGGTTAG	CTCCTTCGGT	CCTCCGATCG	8136
TTGTCAGAAG	TAAGTTGGCC	GCAGTGTTAT	CACTCATGGT	TATGGCAGCA	CTGCATAATT	8196
CTCTTACTGT	CATGCCATCC	GTAAGATGCT	TTTCTGTGAC	TGGTGAGTAC	TCAACCAAGT	8256
CATTCTGAGA	ATAGTGTATG	CGGCGACCGA	GTTGCTCTTG	CCCGGCGTCA	ATACGGGATA	8316
ATACCGCGCC	ACATAGCAGA	ACTTTAAAAG	TGCTCATCAT	TGGAAAACGT	TCTTCGGGGC	8376
GAAAACCTCTC	AAGGATCTTA	CCGCTGTTGA	GATCCAGTTC	GATGTAACCC	ACTCGTGCAC	8436
CCAAC TGATC	TTCAGCATCT	TTTACTTTCA	CCAGCGTTTC	TGGGTGAGCA	AAAACAGGAA	8496
GGCAAAATGC	CGCAAAAAAG	GGAATAAGGG	CGACACGGAA	ATGTTGAATA	CTCATACTCT	8556
TCCTTTTTTCA	ATATTATTGA	AGCATTTATC	AGGGTTATTG	TCTCATGAGC	GGATACATAT	8616
TTGAATGTAT	TTAGAAAAAT	AAACAAATAG	GGGTTCCGCG	CACATTTCCT	CGAAAAGTGC	8676
CACCTGACGT	CTAAGAAACC	ATTATTATCA	TGACATTAAAC	CTATAAAAAT	AGGCGTATCA	8736
CGAGGCCCTT	TCGTCTCGCG	CGTTTCGGTG	ATGACGGTGA	AAACCTCTGA	CACATGCAGC	8796
TCCCGGAGAC	GGTCACAGCT	TGTCTGTAAG	CGGATGCCGG	GAGCAGACAA	GCCCGTCAGG	8856
GCGCGTCAGC	GGGTGTTGGC	GGGTGTCGGG	GCTGGCTTAA	CTATGCGGCA	TCAGAGCAGA	8916
TTGTACTGAG	AGTGCACCAT	ATGCGGTGTG	AAATACCGCA	CAGATGCGTA	AGGAGAAAAT	8976
ACCGCATCAG	GCGCCATTCT	CCATTCAGGC	TGCGCAACTG	TTGGGAAGGG	CGATCGGTGC	9036
GGGCCTCTTC	GCTATTACGC	CAGCTGGCGA	AAGGGGGATG	TGCTGCAAGG	CGATTAAGTT	9096
GGGTAACGCC	AGGGTTTTCC	CAGTCACGAC	GTTGTAAAAC	GACGGCCAGT	GCCAAGCTTG	9156

(2) INFORMATION FOR SEQ ID NO:3:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 12022 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

- (ix) FEATURE:
 (A) NAME/KEY: CDS
 (B) LOCATION: 1006..3294

- (ix) FEATURE:
 (A) NAME/KEY: CDS
 (B) LOCATION: 6153..8234

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GTCGACGGTA TCGATAAGCT TGATATCGAA TTCCTGCAGC CCGGGGGATC CACTAGTACT	60
CGAGACCTAG GAGTTAATTT TAAAAAGCA GTCAAAAGTC CAAGTGGCCC TTGCGAGCAT	120
TTACTCTCTC TGTTTGCTCT GGTTAATAAT CTCAGGAGCA CAAACATTCC TTACTAGTCC	180
TAGAAGTTAA TTTTAAAAA GCAGTCAAAA GTCCAAGTGG CCCTTGCGAG CATTTACTCT	240
CTCTGTTTGC TCTGGTTAAT AATCTCAGGA GCACAAACAT TCCTTACTAG TTCTAGAGCG	300
GCCGCCAGTG TGCTGGAATT CGGCTTTTTT AGGGCTGGAA GCTACCTTTG ACATCATTTT	360
CTCTGCGAAT GCATGTATAA TTTCTACAGA ACCTATTAGA AAGGATCACC CAGCCTCTGC	420
TTTTGTACAA CTTTCCCTTA AAAAAGTGCC AATTCCACTG CTGTTTGGCC CAATAGTGAG	480
AACTTTTTCC TGCTGCCTCT TGGTGCTTTT GCCTATGGCC CCTATTCTGC CTGCTGAAGA	540
CACTCTTGCC AGCATGGACT TAAACCCCTC CAGCTCTGAC AATCCTCTTT CTCTTTTGTT	600
TTACATGAAG GGTCTGGCAG CCAAAGCAAT CACTCAAAGT TCAAACCTTA TCATTTTTTG	660
CTTTGTTTCT CTTGGCCTTG GTTTTGTACA TCAGCTTTGA AAATACCATC CCAGGGTTAA	720
TGCTGGGGTT AATTTATAAC TAAGAGTGCT CTAGTTTTGC AATACAGGAC ATGCTATAAA	780
AATGGAAAGA TGTTGCTTTC TGAGAGATCT CGAGGAAGCT AACAACAAAG AACAACAAAC	840
AACAATCAGG TAAGTATCCT TTTTACAGCA CAACTTAATG AGACAGATAG AAAGTGGTCT	900
TGTAGAAACA GAGTAGTCGC CTGCTTTTCT GCCAGGTGCT GACTTCTCTC CCCTTCTCTT	960
TTTTCTTTT CTCAGGATAA CAAGAACGAA ACAATAACAG CCACC ATG GAA ATA	1014
	Met Glu Ile

GAG CTC TCC ACC TGC TTC TTT CTG TGC CTT TTG CGA TTC TGC TTT AGT	1062
Glu Leu Ser Thr Cys Phe Phe Leu Cys Leu Leu Arg Phe Cys Phe Ser	
5 10 15	
GCC ACC AGA AGA TAC TAC CTG GGT GCA GTG GAA CTG TCA TGG GAC TAT	1110
Ala Thr Arg Arg Tyr Tyr Leu Gly Ala Val Glu Leu Ser Trp Asp Tyr	
20 25 30 35	
ATG CAA AGT GAT CTC GGT GAG CTG CCT GTG GAC GCA AGA TTT CCT CCT	1158
Met Gln Ser Asp Leu Gly Glu Leu Pro Val Asp Ala Arg Phe Pro Pro	
40 45 50	
AGA GTG CCA AAA TCT TTT CCA TTC AAC ACC TCA GTC GTG TAC AAA AAG	1206
Arg Val Pro Lys Ser Phe Pro Phe Asn Thr Ser Val Val Tyr Lys Lys	
55 60 65	
ACT CTG TTT GTA GAA TTC ACG GTT CAC CTT TTC AAC ATC GCT AAG CCA	1254
Thr Leu Phe Val Glu Phe Thr Val His Leu Phe Asn Ile Ala Lys Pro	
70 75 80	
AGG CCA CCC TGG ATG GGT CTG CTA GGT CCT ACC ATC CAG GCT GAG GTT	1302
Arg Pro Pro Trp Met Gly Leu Leu Gly Pro Thr Ile Gln Ala Glu Val	
85 90 95	
TAT GAT ACA GTG GTC ATT ACA CTT AAG AAC ATG GCT TCC CAT CCT GTC	1350
Tyr Asp Thr Val Val Ile Thr Leu Lys Asn Met Ala Ser His Pro Val	
100 105 110 115	
AGT CTT CAT GCT GTT GGT GTA TCC TAC TGG AAA GCT TCT GAG GGA GCT	1398
Ser Leu His Ala Val Gly Val Ser Tyr Trp Lys Ala Ser Glu Gly Ala	
120 125 130	
GAA TAT GAT GAT CAG ACC AGT CAA AGG GAG AAA GAA GAT GAT AAA GTC	1446
Glu Tyr Asp Asp Gln Thr Ser Gln Arg Glu Lys Glu Asp Asp Lys Val	
135 140 145	
TTC CCT GGT GGA AGC CAT ACA TAT GTC TGG CAG GTC CTG AAA GAG AAT	1494
Phe Pro Gly Ser His Thr Tyr Val Trp Gln Val Leu Lys Glu Asn	
150 155 160	
GGT CCA ATG GCC TCT GAC CCA CTG TGC CTT ACC TAC TCA TAT CTT TCT	1542
Gly Pro Met Ala Ser Asp Pro Leu Cys Leu Thr Tyr Ser Tyr Leu Ser	
165 170 175	
CAT GTG GAC CTG GTA AAA GAC TTG AAT TCA GGC CTC ATT GGA GCC CTA	1590
His Val Asp Leu Val Lys Asp Leu Asn Ser Gly Leu Ile Gly Ala Leu	
180 185 190 195	
CTA GTA TGT AGA GAA GGG AGT CTG GCC AAG GAA AAG ACA CAG ACC TTG	1638
Leu Val Cys Arg Glu Gly Ser Leu Ala Lys Glu Lys Thr Gln Thr Leu	
200 205 210	
CAC AAA TTT ATA CTA CTT TTT GCT GTA TTT GAT GAA GGG AAA AGT TGG	1686
His Lys Phe Ile Leu Leu Phe Ala Val Phe Asp Glu Gly Lys Ser Trp	
215 220 225	

CAC TCA GAA ACA AAG AAC TCC TTG ATG CAG GAT AGG GAT GCT GCA TCT	1734
His Ser Glu Thr Lys Asn Ser Leu Met Gln Asp Arg Asp Ala Ala Ser	
230 235 240	
GCT CGG GCC TGG CCT AAA ATG CAC ACA GTC AAT GGT TAT GTA AAC AGG	1782
Ala Arg Ala Trp Pro Lys Met His Thr Val Asn Gly Tyr Val Asn Arg	
245 250 255	
TCT CTG CCA GGT CTG ATT GGA TGC CAC AGG AAA TCA GTC TAT TGG CAT	1830
Ser Leu Pro Gly Leu Ile Gly Cys His Arg Lys Ser Val Tyr Trp His	
260 265 270 275	
GTG ATT GGA ATG GGC ACC ACT CCT GAA GTG CAC TCA ATA TTC CTC GAA	1878
Val Ile Gly Met Gly Thr Thr Pro Glu Val His Ser Ile Phe Leu Glu	
280 285 290	
GGT CAC ACA TTT CTT GTG AGG AAC CAT CGC CAG GCG TCC TTG GAA ATC	1926
Gly His Thr Phe Leu Val Arg Asn His Arg Gln Ala Ser Leu Glu Ile	
295 300 305	
TCG CCA ATA ACT TTC CTT ACT GCT CAA ACA CTC TTG ATG GAC CTT GGA	1974
Ser Pro Ile Thr Phe Leu Thr Ala Gln Thr Leu Leu Met Asp Leu Gly	
310 315 320	
CAG TTT CTA CTG TTT TGT CAT ATC TCT TCC CAC CAA CAT GAT GGC ATG	2022
Gln Phe Leu Leu Phe Cys His Ile Ser Ser His Gln His Asp Gly Met	
325 330 335	
GAA GCT TAT GTC AAA GTA GAC AGC TGT CCA GAG GAA CCC CAA CTA CGA	2070
Glu Ala Tyr Val Lys Val Asp Ser Cys Pro Glu Glu Pro Gln Leu Arg	
340 345 350 355	
ATG AAA AAT AAT GAA GAA GCG GAA GAC TAT GAT GAT GAT CTT ACT GAT	2118
Met Lys Asn Asn Glu Glu Ala Glu Asp Tyr Asp Asp Asp Leu Thr Asp	
360 365 370	
TCT GAA ATG GAT GTG GTC AGG TTT GAT GAT GAC AAC TCT CCT TCC TTT	2166
Ser Glu Met Asp Val Val Arg Phe Asp Asp Asp Asn Ser Pro Ser Phe	
375 380 385	
ATC CAA ATT CGC TCA GTT GCC AAG AAG CAT CCT AAA ACT TGG GTA CAT	2214
Ile Gln Ile Arg Ser Val Ala Lys Lys His Pro Lys Thr Trp Val His	
390 395 400	
TAC ATT GCT GCT GAA GAG GAG GAC TGG GAC TAT GCT CCC TTA GTC CTC	2262
Tyr Ile Ala Ala Glu Glu Glu Asp Trp Asp Tyr Ala Pro Leu Val Leu	
405 410 415	
GCC CCC GAT GAC AGA AGT TAT AAA AGT CAA TAT TTG AAC AAT GGC CCT	2310
Ala Pro Asp Asp Arg Ser Tyr Lys Ser Gln Tyr Leu Asn Asn Gly Pro	
420 425 430 435	
CAG CGG ATT GGT AGG AAG TAC AAA AAA GTC CGA TTT ATG GCA TAC ACA	2358
Gln Arg Ile Gly Arg Lys Tyr Lys Lys Val Arg Phe Met Ala Tyr Thr	
440 445 450	

GAT GAA ACC TTT AAG ACT CGT GAA GCT ATT CAG CAT GAA TCA GGA ATC Asp Glu Thr Phe Lys Thr Arg Glu Ala Ile Gln His Glu Ser Gly Ile 455 460 465	2406
TTG GGA CCT TTA CTT TAT GGG GAA GTT GGA GAC ACA CTG TTG ATT ATA Leu Gly Pro Leu Leu Tyr Gly Glu Val Gly Asp Thr Leu Leu Ile Ile 470 475 480	2454
TTT AAG AAT CAA GCA AGC AGA CCA TAT AAC ATC TAC CCT CAC GGA ATC Phe Lys Asn Gln Ala Ser Arg Pro Tyr Asn Ile Tyr Pro His Gly Ile 485 490 495	2502
ACT GAT GTC CGT CCT TTG TAT TCA AGG AGA TTA CCA AAA GGT GTA AAA Thr Asp Val Arg Pro Leu Tyr Ser Arg Arg Leu Pro Lys Gly Val Lys 500 505 510 515	2550
CAT TTG AAG GAT TTT CCA ATT CTG CCA GGA GAA ATA TTC AAA TAT AAA His Leu Lys Asp Phe Pro Ile Leu Pro Gly Glu Ile Phe Lys Tyr Lys 520 525 530	2598
TGG ACA GTG ACT GTA GAA GAT GGG CCA ACT AAA TCA GAT CCT CGG TGC Trp Thr Val Thr Val Glu Asp Gly Pro Thr Lys Ser Asp Pro Arg Cys 535 540 545	2646
CTG ACC CGC TAT TAC TCT AGT TTC GTT AAT ATG GAG AGA GAT CTA GCT Leu Thr Arg Tyr Tyr Ser Ser Phe Val Asn Met Glu Arg Asp Leu Ala 550 555 560	2694
TCA GGA CTC ATT GGC CCT CTC CTC ATC TGC TAC AAA GAA TCT GTA GAT Ser Gly Leu Ile Gly Pro Leu Leu Ile Cys Tyr Lys Glu Ser Val Asp 565 570 575	2742
CAA AGA GGA AAC CAG ATA ATG TCA GAC AAG AGG AAT GTC ATC CTG TTT Gln Arg Gly Asn Gln Ile Met Ser Asp Lys Arg Asn Val Ile Leu Phe 580 585 590 595	2790
TCT GTA TTT GAT GAG AAC CGA AGC TGG TAC CTC ACA GAG AAT ATA CAA Ser Val Phe Asp Glu Asn Arg Ser Trp Tyr Leu Thr Glu Asn Ile Gln 600 605 610	2838
CGC TTT CTC CCC AAT CCA GCT GGA GTG CAG CTT GAG GAT CCA GAG TTC Arg Phe Leu Pro Asn Pro Ala Gly Val Gln Leu Glu Asp Pro Glu Phe 615 620 625	2886
CAA GCC TCC AAC ATC ATG CAC AGC ATC AAT GGC TAT GTT TTT GAT AGT Gln Ala Ser Asn Ile Met His Ser Ile Asn Gly Tyr Val Phe Asp Ser 630 635 640	2934
TTG CAG TTG TCA GTT TGT TTG CAT GAG GTG GCA TAC TGG TAC ATT CTA Leu Gln Leu Ser Val Cys Leu His Glu Val Ala Tyr Trp Tyr Ile Leu 645 650 655	2982
AGC ATT GGA GCA CAG ACT GAC TTC CTT TCT GTC TTC TTC TCT GGA TAT Ser Ile Gly Ala Gln Thr Asp Phe Leu Ser Val Phe Phe Ser Gly Tyr 660 665 670 675	3030
ACC TTC AAA CAC AAA ATG GTC TAT GAA GAC ACA CTC ACC CTA TTC CCA	3078

Thr	Phe	Lys	His	Lys	Met	Val	Tyr	Glu	Asp	Thr	Leu	Thr	Leu	Phe	Pro	
680								685				690				
TTC	TCA	GGA	GAA	ACT	GTC	TTC	ATG	TCG	ATG	GAA	AAC	CCA	GGT	CTA	TGG	3126
Phe	Ser	Gly	Glu	Thr	Val	Phe	Met	Ser	Met	Glu	Asn	Pro	Gly	Leu	Trp	
695				700				705								
ATT	CTG	GGG	TGC	CAC	AAC	TCA	GAC	TTT	CGG	AAC	AGA	GGC	ATG	ACC	GCC	3174
Ile	Leu	Gly	Cys	His	Asn	Ser	Asp	Phe	Arg	Asn	Arg	Gly	Met	Thr	Ala	
710				715				720								
TTA	CTG	AAG	GTT	TCT	AGT	TGT	GAC	AAG	AAC	ACT	GGT	GAT	TAT	TAC	GAG	3222
Leu	Leu	Lys	Val	Ser	Ser	Cys	Asp	Lys	Asn	Thr	Gly	Asp	Tyr	Tyr	Glu	
725				730				735								
GAC	AGT	TAT	GAA	GAT	ATT	TCA	GCA	TAC	TTG	CTG	AGT	AAA	AAC	AAT	GCC	3270
Asp	Ser	Tyr	Glu	Asp	Ile	Ser	Ala	Tyr	Leu	Leu	Ser	Lys	Asn	Asn	Ala	
740				745				750				755				
ATT	GAA	CCA	AGA	AGC	TTC	TCC	CAG	GTAAGTTATT	ATATAAAATTC	AAGACACCCT						3324
Ile	Glu	Pro	Arg	Ser	Phe	Ser	Gln									
760																
AGCACTAGGC	AAAAGCAATT	TAATGCCACC	ACAATTCCAG	AAAATGACAT	AGAGAAGACT											3384
GACCCTTGGT	TTGCACACAG	AACACCTATG	CCTAAAATAC	AAAATGTCTC	CTCTAGTGAT											3444
TTGTTGATGC	TCTTGCGACA	GAGTCCTACT	CCACATGGGC	TATCCTTATC	TGATCTCCAA											3504
GAAGCCAAAT	ATGAGACTTT	TTCTGATGAT	CCATCACCTG	GAGCAATAGA	CAGTAATAAC											3564
AGCCTGTCTG	AAATGACACA	CTTCAGGCCA	CAGCTCCATC	ACAGTGGGGA	CATGGTATTT											3624
ACCCCTGAGT	CAGGCCTCCA	ATTAAGATTA	AATGAGAAAC	TGGGGACAAC	TGCAGCAACA											3684
GAGTTGAAGA	AACTTGATTT	CAAAGTTTCT	AGTACATCAA	ATAATCTGAT	TTCAACAATT											3744
CCATCAGACA	ATTTGGCAGC	AGGTACTGAT	AATACAAGTT	CCTTAGGACC	CCCAAGTATG											3804
CCAGTTCATT	ATGATAGTCA	ATTAGATACC	ACTCTATTTG	GCAAAAAGTC	ATCTCCCCTT											3864
ACTGAGTCTG	GTGGACCTCT	GAGCTTGAGT	GAAGAAAATA	ATGATTCAAA	GTTGTTAGAA											3924
TCAGGTTTAA	TGAATAGCCA	AGAAAGTTCA	TGGGGAAAAA	ATGTATCGTC	AACAGAGAGT											3984
GGTAGGTTAT	TTAAAGGGAA	AAGAGCTCAT	GGACCTGCTT	TGTTGACTAA	AGATAATGCC											4044
TTATTCAAAG	TTAGCATCTC	TTTGTTAAAG	ACAAACAAAA	CTTCCAATAA	TTCAGCAACT											4104
AATAGAAAGA	CTCACATTGA	TGGCCCATCA	TTATTAATTG	AGAATAGTCC	ATCAGTCTGG											4164
CAAAATATAT	TAGAAAGTGA	CACTGAGTTT	AAAAAAGTGA	CACCTTTGAT	TCATGACAGA											4224
ATGCTTATGG	ACAAAAATGC	TACAGCTTTG	AGGCTAAATC	ATATGTCAAA	TAAAACTACT											4284
TCATCAAAAA	ACATGGAAAT	GGTCCAACAG	AAAAAAGAGG	GCCCCATTCC	ACCAGATGCA											4344

CAAAATCCAG	ATATGTCGTT	CTTTAAGATG	CTATTCTTGC	CAGAATCAGC	AAGGTGGATA	4404
CAAAGGACTC	ATGGAAAGAA	CTCTCTGAAC	TCTGGGCAAG	GCCCCAGTCC	AAAGCAATTA	4464
GTATCCTTAG	GACCAGAAAA	ATCTGTGGAA	GGTCAGAATT	TCTTGTCTGA	GAAAAACAAA	4524
GTGGTAGTAG	GAAAGGGTGA	ATTTACAAAG	GACGTAGGAC	TCAAAGAGAT	GGTTTTTCCA	4584
AGCAGCAGAA	ACCTATTTCT	TACTAACTTG	GATAATTTAC	ATGAAAATAA	TACACACAAT	4644
CAAGAAAAAA	AAATTCAGGA	AGAAATAGAA	AAGAAGGAAA	CATTAATCCA	AGAGAATGTA	4704
GTTTTGCCTC	AGATACATAC	AGTGACTGGC	ACTAAGAATT	TCATGAAGAA	CCTTTTCTTA	4764
CTGAGCACTA	GGCAAAATGT	AGAAGGTTCA	TATGAGGGGG	CATATGCTCC	AGTACTTCAA	4824
GATTTTAGGT	CATTAAATGA	TTCAACAAAT	AGAACAAAGA	AACACACAGC	TCATTTCTCA	4884
AAAAAAGGGG	AGGAAGAAAA	CTTGGAAGGC	TTGGGAAATC	AAACCAAGCA	AATTGTAGAG	4944
AAATATGCAT	GCACCACAAG	GATATCTCCT	AATACAAGCC	AGCAGAATTT	TGTCACGCAA	5004
CGTAGTAAGA	GAGCTTTGAA	ACAATTCAGA	CTCCCACTAG	AAGAAACAGA	ACTTGAAAAA	5064
AGGATAATTG	TGGATGACAC	CTCAACCCAG	TGGTCCAAAA	ACATGAAACA	TTTGACCCCG	5124
AGCACCTCA	CACAGATAGA	CTACAATGAG	AAGGAGAAAG	GGGCCATTAC	TCAGTCTCCC	5184
TTATCAGATT	GCCTTACGAG	GAGTCATAGC	ATCCCTCAAG	CAAATAGATC	TCCATTACCC	5244
ATTGCAAAGG	TATCATCATT	TCCATCTATT	AGACCTATAT	ATCTGACCAG	GGTCCTATTC	5304
CAAGACAACT	CTTCTCATCT	TCCAGCAGCA	TCTTATAGAA	AGAAAGATTC	TGGGGTCCAA	5364
GAAAGCAGTC	ATTTCTTACA	AGGAGCCAAA	AAAAATAACC	TTTCTTTAGC	CATTCTAACC	5424
TTGGAGATGA	CTGGTGATCA	AAGAGAGGTT	GGCTCCCTGG	GGACAAGTGC	CACAAATTCA	5484
GTCACATACA	AGAAAGTTGA	GAACACTGTT	CTCCCGAAAC	CAGACTTGCC	CAAAACATCT	5544
GGCAAAGTTG	AATTGCTTCC	AAAAGTTCAC	ATTTATCAGA	AGGACCTATT	CCCTACGGAA	5604
ACTAGCAATG	GGTCTCCTGG	CCATCTGGAT	CTCGTGGAAG	GGAGCCTTCT	TCAGGGAACA	5664
GAGGGAGCGA	TTAAGTGGAA	TGAAGCAAAC	AGACCTGGAA	AAGTTCCCTT	TCTGAGAGTA	5724
GCAACAGAAA	GCTCTGCAAA	GACTCCCTCC	AAGCTATTGG	ATCCTCTTGC	TTGGGATAAC	5784
CACTATGGTA	CTCAGATACC	AAAAGAAGAG	TGGAAATCCC	AAGAGAAGTC	ACCAGAAAAA	5844
ACAGCTTTTA	AGAAAAAGGA	TACCATTTTG	TCCCTGAACG	CTTGTGAAAG	CAATCATGCA	5904
ATAGCAGCAA	TAAATGAGGG	ACAAAATAAG	CCCGAAATAG	AAGTCACCTG	GGCAAAGCAA	5964
GGTAGGACTG	AAAGGCTGTG	CTCTCAATTG	TGCTAATAAA	GCTTGGCAAG	AGTATTTCAA	6024
GGAAGATGAA	GTCATTAACT	ATGCAAAATG	CTTCTCAGGC	ACCTAGGAAA	ATGAGGATGT	6084

GAGGCATTTC TACCCACTTG GTACATAAAA TTATTGGGTC ACCCTTTTCC TCTTCTTTTT	6144
TTCTCCAG AAC CCA CCA GTC TTG AAA CGC CAT CAA CGG GAA ATA ACT CGT	6194
Asn Pro Pro Val Leu Lys Arg His Gln Arg Glu Ile Thr Arg	
1 5 10	
ACT ACT CTT CAG TCA GAT CAA GAG GAA ATT GAC TAT GAT GAT ACC ATA	6242
Thr Thr Leu Gln Ser Asp Gln Glu Glu Ile Asp Tyr Asp Asp Thr Ile	
15 20 25 30	
TCA GTT GAA ATG AAG AAG GAA GAT TTT GAC ATT TAT GAT GAG GAT GAA	6290
Ser Val Glu Met Lys Lys Glu Asp Phe Asp Ile Tyr Asp Glu Asp Glu	
35 40 45	
AAT CAG AGC CCC CGC AGC TTT CAA AAG AAA ACA CGA CAC TAT TTT ATT	6338
Asn Gln Ser Pro Arg Ser Phe Gln Lys Lys Thr Arg His Tyr Phe Ile	
50 55 60	
GCT GCA GTG GAG AGG CTC TGG GAT TAT GGG ATG AGT AGC TCC CCA CAT	6386
Ala Ala Val Glu Arg Leu Trp Asp Tyr Gly Met Ser Ser Ser Pro His	
65 70 75	
GTT CTA AGA AAC AGG GCT CAG AGT GGC AGT GTC CCT CAG TTC AAG AAA	6434
Val Leu Arg Asn Arg Ala Gln Ser Gly Ser Val Pro Gln Phe Lys Lys	
80 85 90	
GTT GTT TTC CAG GAA TTT ACT GAT GGC TCC TTT ACT CAG CCC TTA TAC	6482
Val Val Phe Gln Glu Phe Thr Asp Gly Ser Phe Thr Gln Pro Leu Tyr	
95 100 105 110	
CGT GGA GAA CTA AAT GAA CAT TTG GGA CTC CTG GGG CCA TAT ATA AGA	6530
Arg Gly Glu Leu Asn Glu His Leu Gly Leu Leu Gly Pro Tyr Ile Arg	
115 120 125	
GCA GAA GTT GAA GAT AAT ATC ATG GTA ACT TTC AGA AAT CAG GCC TCT	6578
Ala Glu Val Glu Asp Asn Ile Met Val Thr Phe Arg Asn Gln Ala Ser	
130 135 140	
CGT CCC TAT TCC TTC TAT TCT AGC CTT ATT TCT TAT GAG GAA GAT CAG	6626
Arg Pro Tyr Ser Phe Tyr Ser Ser Leu Ile Ser Tyr Glu Glu Asp Gln	
145 150 155	
AGG CAA GGA GCA GAA CCT AGA AAA AAC TTT GTC AAG CCT AAT GAA ACC	6674
Arg Gln Gly Ala Glu Pro Arg Lys Asn Phe Val Lys Pro Asn Glu Thr	
160 165 170	
AAA ACT TAC TTT TGG AAA GTG CAA CAT CAT ATG GCA CCC ACT AAA GAT	6722
Lys Thr Tyr Phe Trp Lys Val Gln His His Met Ala Pro Thr Lys Asp	
175 180 185 190	
GAG TTT GAC TGC AAA GCC TGG GCT TAT TTC TCT GAT GTT GAC CTG GAA	6770
Glu Phe Asp Cys Lys Ala Trp Ala Tyr Phe Ser Asp Val Asp Leu Glu	
195 200 205	
AAA GAT GTG CAC TCA GGC CTG ATT GGA CCC CTT CTG GTC TGC CAC ACT	6818
Lys Asp Val His Ser Gly Leu Ile Gly Pro Leu Leu Val Cys His Thr	

210	215	220	
AAC ACA CTG AAC CCT GCT CAT GGG AGA CAA GTG ACA GTA CAG GAA TTT Asn Thr Leu Asn Pro Ala His Gly Arg Gln Val Thr Val Gln Glu Phe 225 230 235			6866
GCT CTG TTT TTC ACC ATC TTT GAT GAG ACC AAA AGC TGG TAC TTC ACT Ala Leu Phe Phe Thr Ile Phe Asp Glu Thr Lys Ser Trp Tyr Phe Thr 240 245 250			6914
GAA AAT ATG GAA AGA AAC TGC AGG GCT CCC TGC AAT ATC CAG ATG GAA Glu Asn Met Glu Arg Asn Cys Arg Ala Pro Cys Asn Ile Gln Met Glu 255 260 265 270			6962
GAT CCC ACT TTT AAA GAG AAT TAT CGC TTC CAT GCA ATC AAT GGC TAC Asp Pro Thr Phe Lys Glu Asn Tyr Arg Phe His Ala Ile Asn Gly Tyr 275 280 285			7010
ATA ATG GAT ACA CTA CCT GGC TTA GTA ATG GCT CAG GAT CAA AGG ATT Ile Met Asp Thr Leu Pro Gly Leu Val Met Ala Gln Asp Gln Arg Ile 290 295 300			7058
CGA TGG TAT CTG CTC AGC ATG GGC AGC AAT GAA AAC ATC CAT TCT ATT Arg Trp Tyr Leu Leu Ser Met Gly Ser Asn Glu Asn Ile His Ser Ile 305 310 315			7106
CAT TTC AGT GGA CAT GTG TTC ACT GTA CGA AAA AAA GAG GAG TAT AAA His Phe Ser Gly His Val Phe Thr Val Arg Lys Lys Glu Glu Tyr Lys 320 325 330			7154
ATG GCA CTG TAC AAT CTC TAT CCA GGT GTT TTT GAG ACA GTG GAA ATG Met Ala Leu Tyr Asn Leu Tyr Pro Gly Val Phe Glu Thr Val Glu Met 335 340 345 350			7202
TTA CCA TCC AAA GCT GGA ATT TGG CGG GTG GAA TGC CTT ATT GGC GAG Leu Pro Ser Lys Ala Gly Ile Trp Arg Val Glu Cys Leu Ile Gly Glu 355 360 365			7250
CAT CTA CAT GCT GGG ATG AGC ACA CTT TTT CTG GTG TAC AGC AAT AAG His Leu His Ala Gly Met Ser Thr Leu Phe Leu Val Tyr Ser Asn Lys 370 375 380			7298
TGT CAG ACT CCC CTG GGA ATG GCT TCT GGA CAC ATT AGA GAT TTT CAG Cys Gln Thr Pro Leu Gly Met Ala Ser Gly His Ile Arg Asp Phe Gln 385 390 395			7346
ATT ACA GCT TCA GGA CAA TAT GGA CAG TGG GCC CCA AAG CTG GCC AGA Ile Thr Ala Ser Gly Gln Tyr Gly Gln Trp Ala Pro Lys Leu Ala Arg 400 405 410			7394
CTT CAT TAT TCC GGA TCA ATC AAT GCC TGG AGC ACC AAG GAG CCC TTT Leu His Tyr Ser Gly Ser Ile Asn Ala Trp Ser Thr Lys Glu Pro Phe 415 420 425 430			7442
TCT TGG ATC AAG GTG GAT CTG TTG GCA CCA ATG ATT ATT CAC GGC ATC Ser Trp Ile Lys Val Asp Leu Leu Ala Pro Met Ile Ile His Gly Ile 435 440 445			7490

AAG ACC CAG GGT GCC CGT CAG AAG TTC TCC AGC CTC TAC ATC TCT CAG Lys Thr Gln Gly Ala Arg Gln Lys Phe Ser Ser Leu Tyr Ile Ser Gln 450 455 460	7538
TTT ATC ATC ATG TAT AGT CTT GAT GGG AAG AAG TGG CAG ACT TAT CGA Phe Ile Ile Met Tyr Ser Leu Asp Gly Lys Lys Trp Gln Thr Tyr Arg 465 470 475	7586
GGA AAT TCC ACT GGA ACC TTA ATG GTC TTC TTT GGC AAT GTG GAT TCA Gly Asn Ser Thr Gly Thr Leu Met Val Phe Phe Gly Asn Val Asp Ser 480 485 490	7634
TCT GGG ATA AAA CAC AAT ATT TTT AAC CCT CCA ATT ATT GCT CGA TAC Ser Gly Ile Lys His Asn Ile Phe Asn Pro Pro Ile Ile Ala Arg Tyr 495 500 505 510	7682
ATC CGT TTG CAC CCA ACT CAT TAT AGC ATT CGC AGC ACT CTT CGC ATG Ile Arg Leu His Pro Thr His Tyr Ser Ile Arg Ser Thr Leu Arg Met 515 520 525	7730
GAG TTG ATG GGC TGT GAT TTA AAT AGT TGC AGC ATG CCA TTG GGA ATG Glu Leu Met Gly Cys Asp Leu Asn Ser Cys Ser Met Pro Leu Gly Met 530 535 540	7778
GAG AGT AAA GCA ATA TCA GAT GCA CAG ATT ACT GCT TCA TCC TAC TTT Glu Ser Lys Ala Ile Ser Asp Ala Gln Ile Thr Ala Ser Ser Tyr Phe 545 550 555	7826
ACC AAT ATG TTT GCC ACC TGG TCT CCT TCA AAA GCT CGA CTT CAC CTC Thr Asn Met Phe Ala Thr Trp Ser Pro Ser Lys Ala Arg Leu His Leu 560 565 570	7874
CAA GGG AGG AGT AAT GCC TGG AGA CCT CAG GTG AAT AAT CCA AAA GAG Gln Gly Arg Ser Asn Ala Trp Arg Pro Gln Val Asn Asn Pro Lys Glu 575 580 585 590	7922
TGG CTG CAA GTG GAC TTC CAG AAG ACA ATG AAA GTC ACA GGA GTA ACT Trp Leu Gln Val Asp Phe Gln Lys Thr Met Lys Val Thr Gly Val Thr 595 600 605	7970
ACT CAG GGA GTA AAA TCT CTG CTT ACC AGC ATG TAT GTG AAG GAG TTC Thr Gln Gly Val Lys Ser Leu Leu Thr Ser Met Tyr Val Lys Glu Phe 610 615 620	8018
CTC ATC TCC AGC AGT CAA GAT GGC CAT CAG TGG ACT CTC TTT TTT CAG Leu Ile Ser Ser Ser Gln Asp Gly His Gln Trp Thr Leu Phe Phe Gln 625 630 635	8066
AAT GGC AAA GTA AAG GTT TTT CAG GGA AAT CAA GAC TCC TTC ACA CCT Asn Gly Lys Val Lys Val Phe Gln Gly Asn Gln Asp Ser Phe Thr Pro 640 645 650	8114
GTG GTG AAC TCT CTA GAC CCA CCG TTA CTG ACT CGC TAC CTT CGA ATT Val Val Asn Ser Leu Asp Pro Pro Leu Leu Thr Arg Tyr Leu Arg Ile 655 660 665 670	8162

CAC CCC CAG AGT TGG GTG CAC CAG ATT GCC CTG AGG ATG GAG GTT CTG	8210
His Pro Gln Ser Trp Val His Gln Ile Ala Leu Arg Met Glu Val Leu	
675 680 685	
GGC TGC GAG GCA CAG GAC CTC TAC TGAGGGTGGC CACTGCAGCA CCTGCCACTG	8264
Gly Cys Glu Ala Gln Asp Leu Tyr	
690	
CCGTCACCTC TCCCTCCTCA GCTCCAGGGC AGTGTCCCTC CCTGGCTTGC CTTCTACCTT	8324
TGTGCTAAAT CCTAGCAGAC ACTGCCTTGA AGCCTCCTGA ATTAACATATC ATCAGTCCTG	8384
CATTTCTTTG GTGGGGGGCC AGGAGGGTGC ATCCAATTTA ACTTAACTCT TACCTATTTT	8444
CTGCAGCTGC TCCCAGATTA CTCCTTCCTT CCAATATAAC TAGGCAAAAA GAAGTGAGGA	8504
GAAACCTGCA TGAAAGCATT CTTCCCTGAA AAGTTAGGCC TCTCAGAGTC ACCACTTCCT	8564
CTGTTGTAGA AAAACTATGT GATGAAACTT TGAAAAAGAT ATTTATGATG TTAACCTGTT	8624
TATTGCAGCT TATAATGGTT ACAAATAAAG CAATAGCATC ACAAATTTCA CAAATAAAGC	8684
ATTTTTTTTCA CTGCATTCTA GTTGTGGTTT GTCCAAACTC ATCAATGTAT CTTATCATGT	8744
CTGGATCCCC GGGTGGCATC CCTGTGACCC CTCCCCAGTG CCTCTCCTGG CCCTGGAAGT	8804
TGCCACTCCA GTGCCCACCA GCCTTGTCTT AATAAAATTA AGTTGCATCA TTTTGTCTGA	8864
CTAGGTGTCC TTCTATAATA TTATGGGGTG GAGGGGGGTG GTATGGAGCA AGGGGCAAGT	8924
TGGGAAGACA ACCTGTAGGG CCTGCGGGGT CTATTCGGGA ACCAAGCTGG AGTGCAGTGG	8984
CACAATCTTG GCTCACTGCA ATCTCCGCCT CCTGGGTTCA AGCGATTCTC CTGCCTCAGC	9044
CTCCCGAGTT GTTGGGATTC CAGGCATGCA TGACCAGGCT CAGCTAATTT TTGTTTTTTT	9104
GGTAGAGACG GGGTTTCACC ATATTGGCCA GGCTGGTCTC CAACTCCTAA TCTCAGGTGA	9164
TCTACCCACC TTGGCCTCCC AAATTGCTGG GATTACAGGC GTGAACCACT GCTCCCTTCC	9224
CTGTCCTTCT GATTTTAAAA TAAC'TATACC AGCAGGAGGA CGTCCAGACA CAGCATAGGC	9284
TACCTGCCAT GCCCAACCGG TGGGACATTT GAGTTGCTTG CTTGGCACTG TCCTCTCATG	9344
CGTTGGGTCC ACTCAGTAGA TGCCTGTTGA ATTCGTAATC ATGGTCATAG CTGTTTCCTG	9404
TGTGAAATTG TTATCCGCTC ACAATTCCAC ACAACATACG AGCCGGAAGC ATAAAGTGTA	9464
AAGCCTGGGG TGCCTAATGA GTGAGCTAAC TCACATTAAT TGCGTTGCGC TCACTGCCCG	9524
CTTTCCAGTC GGGAAACCTG TCGTGCCAGC TGCATTAATG AATCGGCCAA CGCGCGGGGA	9584
GAGGCGGTTT GCGTATTGGG CGCTCTTCCG CTTCTCTGCT CACTGACTCG CTGCGCTCGG	9644
TCGTTCCGCT GCGGCGAGCG GTATCAGCTC ACTCAAAGGC GGTAATACGG TTATCCACAG	9704
AATCAGGGGA TAACGCAGGA AAGAACATGT GAGCAAAAGG CCAGCAAAAG GCCAGGAACC	9764

GTAAAAAGGC CGCGTTGCTG GCGTTTTTTC ATAGGCTCCG CCCCCCTGAC GAGCATCACA 9824
 AAAATCGACG CTCAAGTCAG AGGTGGCGAA ACCCGACAGG ACTATAAAGA TACCAGGCGT 9884
 TTCCCCCTGG AAGCTCCCTC GTGCGCTCTC CTGTTCCGAC CCTGCCGCTT ACCGGATACC 9944
 TGTCCGCCTT TCTCCCTTCG GGAAGCGTGG CGCTTTCTCA TAGCTCACGC TGTAGGTATC 10004
 TCAGTTCGGT GTAGGTCGTT CGCTCCAAGC TGGGCTGTGT GCACGAACCC CCCGTTTCAGC 10064
 CCGACCGCTG CGCCTTATCC GGTAAGTATC GTCTTGAGTC CAACCCGGTA AGACACGACT 10124
 TATCGCCACT GGCAGCAGCC ACTGGTAACA GGATTAGCAG AGCGAGGTAT GTAGGCGGTG 10184
 CTACAGAGTT CTTGAAGTGG TGGCCTAACT ACGGCTACAC TAGAAGGACA GTATTTGGTA 10244
 TCTGCGCTCT GCTGAAGCCA GTTACCTTCG GAAAAAGAGT TGGTAGCTCT TGATCCGGCA 10304
 AACAAACCAC CGCTGGTAGC GGTGGTTTTT TTGTTTGCAA GCAGCAGATT ACGCGCAGAA 10364
 AAAAAGGATC TCAAGAAGAT CCTTTGATCT TTTCTACGGG GTCTGACGCT CAGTGAACG 10424
 AAAACTCACG TTAAGGGATT TTGGTCATGA GATTATCAAA AAGGATCTTC ACCTAGATCC 10484
 TTTTAAATTA AAAATGAAGT TTAAATCAA TCTAAAGTAT ATATGAGTAA ACTTGGTCTG 10544
 ACAGTTACCA ATGCTTAATC AGTGAGGCAC CTATCTCAGC GATCTGTCTA TTTCGTTTCT 10604
 CCATAGTTGC CTGACTCCCC GTCGTGTAGA TAACTACGAT ACGGGAGGGC TTACCATCTG 10664
 GCCCCAGTGC TGCAATGATA CCGCGAGACC CACGCTCACC GGCTCCAGAT TTATCAGCAA 10724
 TAAACCAGCC AGCCGGAAGG GCCGAGCGCA GAAGTGGTCC TGCAACTTTA TCCGCCTCCA 10784
 TCCAGTCTAT TAATTGTTGC CGGGAAGCTA GAGTAAGTAG TTCGCCAGTT AATAGTTTGC 10844
 GCAACGTTGT TGCCATTGCT ACAGGCATCG TGGTGTACAG CTCGTCGTTT GGTATGGCTT 10904
 CATTACGCTC CGGTTCCCAA CGATCAAGGC GAGTTACATG ATCCCCCATG TTGTGCAAAA 10964
 AAGCGGTTAG CTCCTTCGGT CCTCCGATCG TTGTCAGAAG TAAGTTGGCC GCAGTGTTAT 11024
 CACTCATGGT TATGGCAGCA CTGCATAATT CTCTTACTGT CATGCCATCC GTAAGATGCT 11084
 TTTCTGTGAC TGGTGAGTAC TCAACCAAGT CATTCTGAGA ATAGTGATG CGGCGACCGA 11144
 GTTGCTCTTG CCCGGCGTCA ATACGGGATA ATACCGCGCC ACATAGCAGA ACTTTAAAAG 11204
 TGCTCATCAT TGGAAAACGT TCTTCGGGGC GAAAACTCTC AAGGATCTTA CCGCTGTTGA 11264
 GATCCAGTTC GATGTAACCC ACTCGTGCAC CCAACTGATC TTCAGCATCT TTTACTTTCA 11324
 CCAGCGTTTC TGGGTGAGCA AAAACAGGAA GGCAAAATGC CGCAAAAAAG GGAATAAGGG 11384
 CGACACGGAA ATGTTGAATA CTCATACTCT TCCTTTTTCA ATATTATTGA AGCATTTATC 11444

AGGGTTATTG TCTCATGAGC GGATACATAT TTGAATGTAT TTAGAAAAAT AAACAAATAG 11504
 GGGTTCCGCG CACATTTCCC CGAAAAGTGC CACCTGACGT CTAAGAAACC ATTATTATCA 11564
 TGACATTAAC CTATAAAAAT AGGCGTATCA CGAGGCCCTT TCGTCTCGCG CGTTTCGGTG 11624
 ATGACGGTGA AAACCTCTGA CACATGCAGC TCCCGGAGAC GGTACACAGCT TGTCTGTAAG 11684
 CGGATGCCCG GAGCAGACAA GCCCGTCAGG GCGCGTCAGC GGGTGTGGC GGGTGTGGG 11744
 GCTGGCTTAA CTATGCGGCA TCAGAGCAGA TTGTACTGAG AGTGCACCAT ATGCGGTGTG 11804
 AAATACCGCA CAGATGCGTA AGGAGAAAAT ACCGCATCAG GCGCCATTTC CCATTCAGGC 11864
 TCGCAACTG TTGGGAAGGG CGATCGGTGC GGGCCTCTTC GCTATTACGC CAGCTGGCGA 11924
 AAGGGGGATG TGCTGCAAGG CGATTAAGTT GGGTAACGCC AGGGTTTTC CAGTCACGAC 11984
 GTTGTA AAC GACGGCCAGT GCCAAGCTTG GGCTGCAG 12022

(2) INFORMATION FOR SEQ ID NO:4:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 11846 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

- (ix) FEATURE:
 - (A) NAME/KEY: CDS
 - (B) LOCATION: 1006..8058

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GTCGACGGTA TCGATAAGCT TGATATCGAA TTCCTGCAGC CCGGGGGATC CACTAGTACT	60
CGAGACCTAG GAGTTAATTT TAAAAAGCA GTCAAAAGTC CAAGTGGCCC TTGCGAGCAT	120
TTACTCTCTC TGTTTGCTCT GGTAAATAAT CTCAGGAGCA CAAACATTCC TTACTAGTCC	180
TAGAAGTTAA TTTTAAAAA GCAGTCAAAA GTCCAAGTGG CCCTTGCGAG CATTTACTCT	240
CTCTGTTTGC TCTGGTTAAT AATCTCAGGA GCACAAACAT TCCTTACTAG TTCTAGAGCG	300
GCCGCCAGTG TGCTGGAATT CGGCTTTTTT AGGGCTGGAA GCTACCTTTG ACATCATTTT	360
CTCTGCGAAT GCATGTATAA TTTCTACAGA ACCTATTAGA AAGGATCACC CAGCCTCTGC	420
TTTTGTACAA CTTTCCCTTA AAAAAGTGGC AATTCCACTG CTGTTTGGCC CAATAGTGAG	480
AACTTTTTC TGCTGCCTCT TGGTGCTTTT GCCTATGGCC CCTATTCTGC CTGCTGAAGA	540
CACTCTTGCC AGCATGGACT TAAACCCCTC CAGCTCTGAC AATCCTCTTT CTCTTTTGT	600
TTACATGAAG GGTCTGGCAG CCAAAGCAAT CACTCAAAGT TCAAACCTTA TCATTTTTTG	660

CTTTGTTCTT	CTTGGCCTTG	GTTTTGTACA	TCAGCTTTGA	AAATACCATC	CCAGGGTTAA	720
TGCTGGGGTT	AATTTATAAC	TAAGAGTGCT	CTAGTTTTGC	AATACAGGAC	ATGCTATAAA	780
AATGGAAAGA	TGTTGCTTTC	TGAGAGATCT	CGAGGAAGCT	AACAACAAAG	AACAACAAAC	840
AACAATCAGG	TAAGTATCCT	TTTTACAGCA	CAACTTAATG	AGACAGATAG	AAACTGGTCT	900
TGTAGAAACA	GAGTAGTCGC	CTGCTTTTCT	GCCAGGTGCT	GACTTCTCTC	CCCTTCTCTT	960
TTTTCCTTTT	CTCAGGATAA	CAAGAACGAA	ACAATAACAG	CCACC	ATG GAA ATA	1014
					Met Glu Ile	
					1	
GAG CTC TCC ACC TGC TTC TTT CTG TGC CTT TTG CGA TTC TGC TTT AGT	1062					
Glu Leu Ser Thr Cys Phe Phe Leu Cys Leu Leu Arg Phe Cys Phe Ser						
5 10 15						
GCC ACC AGA AGA TAC TAC CTG GGT GCA GTG GAA CTG TCA TGG GAC TAT	1110					
Ala Thr Arg Arg Tyr Tyr Leu Gly Ala Val Glu Leu Ser Trp Asp Tyr						
20 25 30 35						
ATG CAA AGT GAT CTC GGT GAG CTG CCT GTG GAC GCA AGA TTT CCT CCT	1158					
Met Gln Ser Asp Leu Gly Glu Leu Pro Val Asp Ala Arg Phe Pro Pro						
40 45 50						
AGA GTG CCA AAA TCT TTT CCA TTC AAC ACC TCA GTC GTG TAC AAA AAG	1206					
Arg Val Pro Lys Ser Phe Pro Phe Asn Thr Ser Val Val Tyr Lys Lys						
55 60 65						
ACT CTG TTT GTA GAA TTC ACG GTT CAC CTT TTC AAC ATC GCT AAG CCA	1254					
Thr Leu Phe Val Glu Phe Thr Val His Leu Phe Asn Ile Ala Lys Pro						
70 75 80						
AGG CCA CCC TGG ATG GGT CTG CTA GGT CCT ACC ATC CAG GCT GAG GTT	1302					
Arg Pro Pro Trp Met Gly Leu Leu Gly Pro Thr Ile Gln Ala Glu Val						
85 90 95						
TAT GAT ACA GTG GTC ATT ACA CTT AAG AAC ATG GCT TCC CAT CCT GTC	1350					
Tyr Asp Thr Val Val Ile Thr Leu Lys Asn Met Ala Ser His Pro Val						
100 105 110 115						
AGT CTT CAT GCT GTT GGT GTA TCC TAC TGG AAA GCT TCT GAG GGA GCT	1398					
Ser Leu His Ala Val Gly Val Ser Tyr Trp Lys Ala Ser Glu Gly Ala						
120 125 130						
GAA TAT GAT GAT CAG ACC AGT CAA AGG GAG AAA GAA GAT GAT AAA GTC	1446					
Glu Tyr Asp Asp Gln Thr Ser Gln Arg Glu Lys Glu Asp Asp Lys Val						
135 140 145						
TTC CCT GGT GGA AGC CAT ACA TAT GTC TGG CAG GTC CTG AAA GAG AAT	1494					
Phe Pro Gly Gly Ser His Thr Tyr Val Trp Gln Val Leu Lys Glu Asn						
150 155 160						
GGT CCA ATG GCC TCT GAC CCA CTG TGC CTT ACC TAC TCA TAT CTT TCT	1542					
Gly Pro Met Ala Ser Asp Pro Leu Cys Leu Thr Tyr Ser Tyr Leu Ser						

165	170	175	
CAT GTG GAC CTG GTA AAA GAC TTG AAT TCA GGC CTC ATT GGA GCC CTA His Val Asp Leu Val Lys Asp Leu Asn Ser Gly Leu Ile Gly Ala Leu 180 185 190 195			1590
CTA GTA TGT AGA GAA GGG AGT CTG GCC AAG GAA AAG ACA CAG ACC TTG Leu Val Cys Arg Glu Gly Ser Leu Ala Lys Glu Lys Thr Gln Thr Leu 200 205 210			1638
CAC AAA TTT ATA CTA CTT TTT GCT GTA TTT GAT GAA GGG AAA AGT TGG His Lys Phe Ile Leu Leu Phe Ala Val Phe Asp Glu Gly Lys Ser Trp 215 220 225			1686
CAC TCA GAA ACA AAG AAC TCC TTG ATG CAG GAT AGG GAT GCT GCA TCT His Ser Glu Thr Lys Asn Ser Leu Met Gln Asp Arg Asp Ala Ala Ser 230 235 240			1734
GCT CGG GCC TGG CCT AAA ATG CAC ACA GTC AAT GGT TAT GTA AAC AGG Ala Arg Ala Trp Pro Lys Met His Thr Val Asn Gly Tyr Val Asn Arg 245 250 255			1782
TCT CTG CCA GGT CTG ATT GGA TGC CAC AGG AAA TCA GTC TAT TGG CAT Ser Leu Pro Gly Leu Ile Gly Cys His Arg Lys Ser Val Tyr Trp His 260 265 270 275			1830
GTG ATT GGA ATG GGC ACC ACT CCT GAA GTG CAC TCA ATA TTC CTC GAA Val Ile Gly Met Gly Thr Thr Pro Glu Val His Ser Ile Phe Leu Glu 280 285 290			1878
GGT CAC ACA TTT CTT GTG AGG AAC CAT CGC CAG GCG TCC TTG GAA ATC Gly His Thr Phe Leu Val Arg Asn His Arg Gln Ala Ser Leu Glu Ile 295 300 305			1926
TCG CCA ATA ACT TTC CTT ACT GCT CAA ACA CTC TTG ATG GAC CTT GGA Ser Pro Ile Thr Phe Leu Thr Ala Gln Thr Leu Leu Met Asp Leu Gly 310 315 320			1974
CAG TTT CTA CTG TTT TGT CAT ATC TCT TCC CAC CAA CAT GAT GGC ATG Gln Phe Leu Leu Phe Cys His Ile Ser Ser His Gln His Asp Gly Met 325 330 335			2022
GAA GCT TAT GTC AAA GTA GAC AGC TGT CCA GAG GAA CCC CAA CTA CGA Glu Ala Tyr Val Lys Val Asp Ser Cys Pro Glu Glu Pro Gln Leu Arg 340 345 350 355			2070
ATG AAA AAT AAT GAA GAA GCG GAA GAC TAT GAT GAT GAT CTT ACT GAT Met Lys Asn Asn Glu Glu Ala Glu Asp Tyr Asp Asp Asp Leu Thr Asp 360 365 370			2118
TCT GAA ATG GAT GTG GTC AGG TTT GAT GAT GAC AAC TCT CCT TCC TTT Ser Glu Met Asp Val Val Arg Phe Asp Asp Asp Asn Ser Pro Ser Phe 375 380 385			2166
ATC CAA ATT CGC TCA GTT GCC AAG AAG CAT CCT AAA ACT TGG GTA CAT Ile Gln Ile Arg Ser Val Ala Lys Lys His Pro Lys Thr Trp Val His 390 395 400			2214

TAC ATT GCT GCT GAA GAG GAG GAC TGG GAC TAT GCT CCC TTA GTC CTC	2262
Tyr Ile Ala Ala Glu Glu Asp Trp Asp Tyr Ala Pro Leu Val Leu	
405 410 415	
GCC CCC GAT GAC AGA AGT TAT AAA AGT CAA TAT TTG AAC AAT GGC CCT	2310
Ala Pro Asp Asp Arg Ser Tyr Lys Ser Gln Tyr Leu Asn Asn Gly Pro	
420 425 430 435	
CAG CGG ATT GGT AGG AAG TAC AAA AAA GTC CGA TTT ATG GCA TAC ACA	2358
Gln Arg Ile Gly Arg Lys Tyr Lys Lys Val Arg Phe Met Ala Tyr Thr	
440 445 450	
GAT GAA ACC TTT AAG ACT CGT GAA GCT ATT CAG CAT GAA TCA GGA ATC	2406
Asp Glu Thr Phe Lys Thr Arg Glu Ala Ile Gln His Glu Ser Gly Ile	
455 460 465	
TTG GGA CCT TTA CTT TAT GGG GAA GTT GGA GAC ACA CTG TTG ATT ATA	2454
Leu Gly Pro Leu Leu Tyr Gly Glu Val Gly Asp Thr Leu Leu Ile Ile	
470 475 480	
TTT AAG AAT CAA GCA AGC AGA CCA TAT AAC ATC TAC CCT CAC GGA ATC	2502
Phe Lys Asn Gln Ala Ser Arg Pro Tyr Asn Ile Tyr Pro His Gly Ile	
485 490 495	
ACT GAT GTC CGT CCT TTG TAT TCA AGG AGA TTA CCA AAA GGT GTA AAA	2550
Thr Asp Val Arg Pro Leu Tyr Ser Arg Arg Leu Pro Lys Gly Val Lys	
500 505 510 515	
CAT TTG AAG GAT TTT CCA ATT CTG CCA GGA GAA ATA TTC AAA TAT AAA	2598
His Leu Lys Asp Phe Pro Ile Leu Pro Gly Glu Ile Phe Lys Tyr Lys	
520 525 530	
TGG ACA GTG ACT GTA GAA GAT GGG CCA ACT AAA TCA GAT CCT CGG TGC	2646
Trp Thr Val Thr Val Glu Asp Gly Pro Thr Lys Ser Asp Pro Arg Cys	
535 540 545	
CTG ACC CGC TAT TAC TCT AGT TTC GTT AAT ATG GAG AGA GAT CTA GCT	2694
Leu Thr Arg Tyr Tyr Ser Ser Phe Val Asn Met Glu Arg Asp Leu Ala	
550 555 560	
TCA GGA CTC ATT GGC CCT CTC CTC ATC TGC TAC AAA GAA TCT GTA GAT	2742
Ser Gly Leu Ile Gly Pro Leu Leu Ile Cys Tyr Lys Glu Ser Val Asp	
565 570 575	
CAA AGA GGA AAC CAG ATA ATG TCA GAC AAG AGG AAT GTC ATC CTG TTT	2790
Gln Arg Gly Asn Gln Ile Met Ser Asp Lys Arg Asn Val Ile Leu Phe	
580 585 590 595	
TCT GTA TTT GAT GAG AAC CGA AGC TGG TAC CTC ACA GAG AAT ATA CAA	2838
Ser Val Phe Asp Glu Asn Arg Ser Trp Tyr Leu Thr Glu Asn Ile Gln	
600 605 610	
CGC TTT CTC CCC AAT CCA GCT GGA GTG CAG CTT GAG GAT CCA GAG TTC	2886
Arg Phe Leu Pro Asn Pro Ala Gly Val Gln Leu Glu Asp Pro Glu Phe	
615 620 625	

CAA GCC TCC AAC ATC ATG CAC AGC ATC AAT GGC TAT GTT TTT GAT AGT	2934
Gln Ala Ser Asn Ile Met His Ser Ile Asn Gly Tyr Val Phe Asp Ser	
630 635 640	
TTG CAG TTG TCA GTT TGT TTG CAT GAG GTG GCA TAC TGG TAC ATT CTA	2982
Leu Gln Leu Ser Val Cys Leu His Glu Val Ala Tyr Trp Tyr Ile Leu	
645 650 655	
AGC ATT GGA GCA CAG ACT GAC TTC CTT TCT GTC TTC TTC TCT GGA TAT	3030
Ser Ile Gly Ala Gln Thr Asp Phe Leu Ser Val Phe Phe Ser Gly Tyr	
660 665 670 675	
ACC TTC AAA CAC AAA ATG GTC TAT GAA GAC ACA CTC ACC CTA TTC CCA	3076
Thr Phe Lys His Lys Met Val Tyr Glu Asp Thr Leu Thr Leu Phe Pro	
680 685 690	
TTC TCA GGA GAA ACT GTC TTC ATG TCG ATG GAA AAC CCA GGT CTA TGG	3126
Phe Ser Gly Glu Thr Val Phe Met Ser Met Glu Asn Pro Gly Leu Trp	
695 700 705	
ATT CTG GGG TGC CAC AAC TCA GAC TTT CGG AAC AGA GGC ATG ACC GCC	3174
Ile Leu Gly Cys His Asn Ser Asp Phe Arg Asn Arg Gly Met Thr Ala	
710 715 720	
TTA CTG AAG GTT TCT AGT TGT GAC AAG AAC ACT GGT GAT TAT TAC GAG	3222
Leu Leu Lys Val Ser Ser Cys Asp Lys Asn Thr Gly Asp Tyr Tyr Glu	
725 730 735	
GAC AGT TAT GAA GAT ATT TCA GCA TAC TTG CTG AGT AAA AAC AAT GCC	3270
Asp Ser Tyr Glu Asp Ile Ser Ala Tyr Leu Leu Ser Lys Asn Asn Ala	
740 745 750 755	
ATT GAA CCA AGA AGC TTC TCC CAG AAT TCA AGA CAC CCT AGC ACT AGG	3318
Ile Glu Pro Arg Ser Phe Ser Gln Asn Ser Arg His Pro Ser Thr Arg	
760 765 770	
CAA AAG CAA TTT AAT GCC ACC ACA ATT CCA GAA AAT GAC ATA GAG AAG	3366
Gln Lys Gln Phe Asn Ala Thr Thr Ile Pro Glu Asn Asp Ile Glu Lys	
775 780 785	
ACT GAC CCT TGG TTT GCA CAC AGA ACA CCT ATG CCT AAA ATA CAA AAT	3414
Thr Asp Pro Trp Phe Ala His Arg Thr Pro Met Pro Lys Ile Gln Asn	
790 795 800	
GTC TCC TCT AGT GAT TTG TTG ATG CTC TTG CGA CAG AGT CCT ACT CCA	3462
Val Ser Ser Ser Asp Leu Leu Met Leu Leu Arg Gln Ser Pro Thr Pro	
805 810 815	
CAT GGG CTA TCC TTA TCT GAT CTC CAA GAA GCC AAA TAT GAG ACT TTT	3510
His Gly Leu Ser Leu Ser Asp Leu Gln Glu Ala Lys Tyr Glu Thr Phe	
820 825 830 835	
TCT GAT GAT CCA TCA CCT GGA GCA ATA GAC AGT AAT AAC AGC CTG TCT	3558
Ser Asp Asp Pro Ser Pro Gly Ala Ile Asp Ser Asn Asn Ser Leu Ser	
840 845 850	
GAA ATG ACA CAC TTC AGG CCA CAG CTC CAT CAC AGT GGG GAC ATG GTA	3606

Glu Met Thr His Phe Arg Pro Gln Leu His His Ser Gly Asp Met Val	
855	860 865
TTT ACC CCT GAG TCA GGC CTC CAA TTA AGA TTA AAT GAG AAA CTG GGG	3654
Phe Thr Pro Glu Ser Gly Leu Gln Leu Arg Leu Asn Glu Lys Leu Gly	
870	875 880
ACA ACT GCA GCA ACA GAG TTG AAG AAA CTT GAT TTC AAA GTT TCT AGT	3702
Thr Thr Ala Ala Thr Glu Leu Lys Lys Leu Asp Phe Lys Val Ser Ser	
885	890 895
ACA TCA AAT AAT CTG ATT TCA ACA ATT CCA TCA GAC AAT TTG GCA GCA	3750
Thr Ser Asn Asn Leu Ile Ser Thr Ile Pro Ser Asp Asn Leu Ala Ala	
900	905 910 915
GGT ACT GAT AAT ACA AGT TCC TTA GGA CCC CCA AGT ATG CCA GTT CAT	3798
Gly Thr Asp Asn Thr Ser Ser Leu Gly Pro Pro Ser Met Pro Val His	
920	925 930
TAT GAT AGT CAA TTA GAT ACC ACT CTA TTT GGC AAA AAG TCA TCT CCC	3846
Tyr Asp Ser Gln Leu Asp Thr Thr Leu Phe Gly Lys Lys Ser Ser Pro	
935	940 945
CTT ACT GAG TCT GGT GGA CCT CTG AGC TTG AGT GAA GAA AAT AAT GAT	3894
Leu Thr Glu Ser Gly Gly Pro Leu Ser Leu Ser Glu Glu Asn Asn Asp	
950	955 960
TCA AAG TTG TTA GAA TCA GGT TTA ATG AAT AGC CAA GAA AGT TCA TGG	3942
Ser Lys Leu Leu Glu Ser Gly Leu Met Asn Ser Gln Glu Ser Ser Trp	
965	970 975
GGA AAA AAT GTA TCG TCA ACA GAG AGT GGT AGG TTA TTT AAA GGG AAA	3990
Gly Lys Asn Val Ser Ser Thr Glu Ser Gly Arg Leu Phe Lys Gly Lys	
980	985 990 995
AGA GCT CAT GGA CCT GCT TTG TTG ACT AAA GAT AAT GCC TTA TTC AAA	4038
Arg Ala His Gly Pro Ala Leu Leu Thr Lys Asp Asn Ala Leu Phe Lys	
1000	1005 1010
GTT AGC ATC TCT TTG TTA AAG ACA AAC AAA ACT TCC AAT AAT TCA GCA	4086
Val Ser Ile Ser Leu Leu Lys Thr Asn Lys Thr Ser Asn Asn Ser Ala	
1015	1020 1025
ACT AAT AGA AAG ACT CAC ATT GAT GGC CCA TCA TTA TTA ATT GAG AAT	4134
Thr Asn Arg Lys Thr His Ile Asp Gly Pro Ser Leu Leu Ile Glu Asn	
1030	1035 1040
AGT CCA TCA GTC TGG CAA AAT ATA TTA GAA AGT GAC ACT GAG TTT AAA	4182
Ser Pro Ser Val Trp Gln Asn Ile Leu Glu Ser Asp Thr Glu Phe Lys	
1045	1050 1055
AAA GTG ACA CCT TTG ATT CAT GAC AGA ATG CTT ATG GAC AAA AAT GCT	4230
Lys Val Thr Pro Leu Ile His Asp Arg Met Leu Met Asp Lys Asn Ala	
1060	1065 1070 1075
ACA GCT TTG AGG CTA AAT CAT ATG TCA AAT AAA ACT ACT TCA TCA AAA	4278
Thr Ala Leu Arg Leu Asn His Met Ser Asn Lys Thr Thr Ser Ser Lys	

1080	1085	1090	
AAC ATG GAA ATG GTC CAA CAG AAA AAA GAG GGC CCC ATT CCA CCA GAT Asn Met Glu Met Val Gln Gln Lys Lys Glu Gly Pro Ile Pro Pro Asp 1095 1100 1105			4326
GCA CAA AAT CCA GAT ATG TCG TTC TTT AAG ATG CTA TTC TTG CCA GAA Ala Gln Asn Pro Asp Met Ser Phe Phe Lys Met Leu Phe Leu Pro Glu 1110 1115 1120			4374
TCA GCA AGG TGG ATA CAA AGG ACT CAT GGA AAG AAC TCT CTG AAC TCT Ser Ala Arg Trp Ile Gln Arg Thr His Gly Lys Asn Ser Leu Asn Ser 1125 1130 1135			4422
GGG CAA GGC CCC AGT CCA AAG CAA TTA GTA TCC TTA GGA CCA GAA AAA Gly Gln Gly Pro Ser Pro Lys Gln Leu Val Ser Leu Gly Pro Glu Lys 1140 1145 1150 1155			4470
TCT GTG GAA GGT CAG AAT TTC TTG TCT GAG AAA AAC AAA GTG GTA GTA Ser Val Glu Gly Gln Asn Phe Leu Ser Glu Lys Asn Lys Val Val Val 1160 1165 1170			4518
GGA AAG GGT GAA TTT ACA AAG GAC GTA GGA CTC AAA GAG ATG GTT TTT Gly Lys Gly Glu Phe Thr Lys Asp Val Gly Leu Lys Glu Met Val Phe 1175 1180 1185			4566
CCA AGC AGC AGA AAC CTA TTT CTT ACT AAC TTG GAT AAT TTA CAT GAA Pro Ser Ser Arg Asn Leu Phe Leu Thr Asn Leu Asp Asn Leu His Glu 1190 1195 1200			4614
AAT AAT ACA CAC AAT CAA GAA AAA AAA ATT CAG GAA GAA ATA GAA AAG Asn Asn Thr His Asn Gln Glu Lys Lys Ile Gln Glu Glu Ile Glu Lys 1205 1210 1215			4662
AAG GAA ACA TTA ATC CAA GAG AAT GTA GTT TTG CCT CAG ATA CAT ACA Lys Glu Thr Leu Ile Gln Glu Asn Val Val Leu Pro Gln Ile His Thr 1220 1225 1230 1235			4710
GTG ACT GGC ACT AAG AAT TTC ATG AAG AAC CTT TTC TTA CTG AGC ACT Val Thr Gly Thr Lys Asn Phe Met Lys Asn Leu Phe Leu Leu Ser Thr 1240 1245 1250			4758
AGG CAA AAT GTA GAA GGT TCA TAT GAG GGG GCA TAT GCT CCA GTA CTT Arg Gln Asn Val Glu Gly Ser Tyr Glu Gly Ala Tyr Ala Pro Val Leu 1255 1260 1265			4806
CAA GAT TTT AGG TCA TTA AAT GAT TCA ACA AAT AGA ACA AAG AAA CAC Gln Asp Phe Arg Ser Leu Asn Asp Ser Thr Asn Arg Thr Lys Lys His 1270 1275 1280			4854
ACA GCT CAT TTC TCA AAA AAA GGG GAG GAA GAA AAC TTG GAA GGC TTG Thr Ala His Phe Ser Lys Lys Gly Glu Glu Glu Asn Leu Glu Gly Leu 1285 1290 1295			4902
GGA AAT CAA ACC AAG CAA ATT GTA GAG AAA TAT GCA TGC ACC ACA AGG Gly Asn Gln Thr Lys Gln Ile Val Glu Lys Tyr Ala Cys Thr Thr Arg 1300 1305 1310 1315			4950

ATA TCT CCT AAT ACA AGC CAG CAG AAT TTT GTC ACG CAA CGT AGT AAG	4998
Ile Ser Pro Asn Thr Ser Gln Gln Asn Phe Val Thr Gln Arg Ser Lys	
1320 1325 1330	
AGA GCT TTG AAA CAA TTC AGA CTC CCA CTA GAA GAA ACA GAA CTT GAA	5046
Arg Ala Leu Lys Gln Phe Arg Leu Pro Leu Glu Glu Thr Glu Leu Glu	
1335 1340 1345	
AAA AGG ATA ATT GTG GAT GAC ACC TCA ACC CAG TGG TCC AAA AAC ATG	5094
Lys Arg Ile Ile Val Asp Asp Thr Ser Thr Gln Trp Ser Lys Asn Met	
1350 1355 1360	
AAA CAT TTG ACC CCG AGC ACC CTC ACA CAG ATA GAC TAC AAT GAG AAG	5142
Lys His Leu Thr Pro Ser Thr Leu Thr Gln Ile Asp Tyr Asn Glu Lys	
1365 1370 1375	
GAG AAA GGG GCC ATT ACT CAG TCT CCC TTA TCA GAT TGC CTT ACG AGG	5190
Glu Lys Gly Ala Ile Thr Gln Ser Pro Leu Ser Asp Cys Leu Thr Arg	
1380 1385 1390 1395	
AGT CAT AGC ATC CCT CAA GCA AAT AGA TCT CCA TTA CCC ATT GCA AAG	5238
Ser His Ser Ile Pro Gln Ala Asn Arg Ser Pro Leu Pro Ile Ala Lys	
1400 1405 1410	
GTA TCA TCA TTT CCA TCT ATT AGA CCT ATA TAT CTG ACC AGG GTC CTA	5286
Val Ser Ser Phe Pro Ser Ile Arg Pro Ile Tyr Leu Thr Arg Val Leu	
1415 1420 1425	
TTC CAA GAC AAC TCT TCT CAT CTT CCA GCA GCA TCT TAT AGA AAG AAA	5334
Phe Gln Asp Asn Ser Ser His Leu Pro Ala Ala Ser Tyr Arg Lys Lys	
1430 1435 1440	
GAT TCT GGG GTC CAA GAA AGC AGT CAT TTC TTA CAA GGA GCC AAA AAA	5382
Asp Ser Gly Val Gln Glu Ser Ser His Phe Leu Gln Gly Ala Lys Lys	
1445 1450 1455	
AAT AAC CTT TCT TTA GCC ATT CTA ACC TTG GAG ATG ACT GGT GAT CAA	5430
Asn Asn Leu Ser Leu Ala Ile Leu Thr Leu Glu Met Thr Gly Asp Gln	
1460 1465 1470 1475	
AGA GAG GTT GGC TCC CTG GGG ACA AGT GCC ACA AAT TCA GTC ACA TAC	5478
Arg Glu Val Gly Ser Leu Gly Thr Ser Ala Thr Asn Ser Val Thr Tyr	
1480 1485 1490	
AAG AAA GTT GAG AAC ACT GTT CTC CCG AAA CCA GAC TTG CCC AAA ACA	5526
Lys Lys Val Glu Asn Thr Val Leu Pro Lys Pro Asp Leu Pro Lys Thr	
1495 1500 1505	
TCT GGC AAA GTT GAA TTG CTT CCA AAA GTT CAC ATT TAT CAG AAG GAC	5574
Ser Gly Lys Val Glu Leu Leu Pro Lys Val His Ile Tyr Gln Lys Asp	
1510 1515 1520	
CTA TTC CCT ACG GAA ACT AGC AAT GGG TCT CCT GGC CAT CTG GAT CTC	5622
Leu Phe Pro Thr Glu Thr Ser Asn Gly Ser Pro Gly His Leu Asp Leu	
1525 1530 1535	

GTG GAA GGG AGC CTT CTT CAG GGA ACA GAG GGA GCG ATT AAG TGG AAT	5670
Val Glu Gly Ser Leu Leu Gln Gly Thr Glu Gly Ala Ile Lys Trp Asn	
1540 1545 1550 1555	
GAA GCA AAC AGA CCT GGA AAA GTT CCC TTT CTG AGA GTA GCA ACA GAA	5718
Glu Ala Asn Arg Pro Gly Lys Val Pro Phe Leu Arg Val Ala Thr Glu	
1560 1565 1570	
AGC TCT GCA AAG ACT CCC TCC AAG CTA TTG GAT CCT CTT GCT TGG GAT	5766
Ser Ser Ala Lys Thr Pro Ser Lys Leu Leu Asp Pro Leu Ala Trp Asp	
1575 1580 1585	
AAC CAC TAT GGT ACT CAG ATA CCA AAA GAA GAG TGG AAA TCC CAA GAG	5814
Asn His Tyr Gly Thr Gln Ile Pro Lys Glu Glu Trp Lys Ser Gln Glu	
1590 1595 1600	
AAG TCA CCA GAA AAA ACA GCT TTT AAG AAA AAG GAT ACC ATT TTG TCC	5862
Lys Ser Pro Glu Lys Thr Ala Phe Lys Lys Lys Asp Thr Ile Leu Ser	
1605 1610 1615	
CTG AAC GCT TGT GAA AGC AAT CAT GCA ATA GCA GCA ATA AAT GAG GGA	5910
Leu Asn Ala Cys Glu Ser Asn His Ala Ile Ala Ala Ile Asn Glu Gly	
1620 1625 1630 1635	
CAA AAT AAG CCC GAA ATA GAA GTC ACC TGG GCA AAG CAA GGT AGG ACT	5958
Gln Asn Lys Pro Glu Ile Glu Val Thr Trp Ala Lys Gln Gly Arg Thr	
1640 1645 1650	
GAA AGG CTG TGC TCT CAA AAC CCA CCA GTC TTG AAA CGC CAT CAA CGG	6006
Glu Arg Leu Cys Ser Gln Asn Pro Pro Val Leu Lys Arg His Gln Arg	
1655 1660 1665	
GAA ATA ACT CGT ACT ACT CTT CAG TCA GAT CAA GAG GAA ATT GAC TAT	6054
Glu Ile Thr Arg Thr Thr Leu Gln Ser Asp Gln Glu Glu Ile Asp Tyr	
1670 1675 1680	
GAT GAT ACC ATA TCA GTT GAA ATG AAG AAG GAA GAT TTT GAC ATT TAT	6102
Asp Asp Thr Ile Ser Val Glu Met Lys Lys Glu Asp Phe Asp Ile Tyr	
1685 1690 1695	
GAT GAG GAT GAA AAT CAG AGC CCC CGC AGC TTT CAA AAG AAA ACA CGA	6150
Asp Glu Asp Glu Asn Gln Ser Pro Arg Ser Phe Gln Lys Lys Thr Arg	
1700 1705 1710 1715	
CAC TAT TTT ATT GCT GCA GTG GAG AGG CTC TGG GAT TAT GGG ATG AGT	6198
His Tyr Phe Ile Ala Ala Val Glu Arg Leu Trp Asp Tyr Gly Met Ser	
1720 1725 1730	
AGC TCC CCA CAT GTT CTA AGA AAC AGG GCT CAG AGT GGC AGT GTC CCT	6246
Ser Ser Pro His Val Leu Arg Asn Arg Ala Gln Ser Gly Ser Val Pro	
1735 1740 1745	
CAG TTC AAG AAA GTT GTT TTC CAG GAA TTT ACT GAT GGC TCC TTT ACT	6294
Gln Phe Lys Lys Val Val Phe Gln Glu Phe Thr Asp Gly Ser Phe Thr	
1750 1755 1760	
CAG CCC TTA TAC CGT GGA GAA CTA AAT GAA CAT TTG GGA CTC CTG GGG	6342

Gln Pro Leu Tyr Arg Gly Glu Leu Asn Glu His Leu Gly Leu Leu Gly	
1765	1770 1775
CCA TAT ATA AGA GCA GAA GTT GAA GAT AAT ATC ATG GTA ACT TTC AGA	6390
Pro Tyr Ile Arg Ala Glu Val Glu Asp Asn Ile Met Val Thr Phe Arg	
1780	1785 1790 1795
AAT CAG GCC TCT CGT CCC TAT TCC TTC TAT TCT AGC CTT ATT TCT TAT	6438
Asn Gln Ala Ser Arg Pro Tyr Ser Phe Tyr Ser Ser Leu Ile Ser Tyr	
	1800 1805 1810
GAG GAA GAT CAG AGG CAA GGA GCA GAA CCT AGA AAA AAC TTT GTC AAG	6486
Glu Glu Asp Gln Arg Gln Gly Ala Glu Pro Arg Lys Asn Phe Val Lys	
	1815 1820 1825
CCT AAT GAA ACC AAA ACT TAC TTT TGG AAA GTG CAA CAT CAT ATG GCA	6534
Pro Asn Glu Thr Lys Thr Tyr Phe Trp Lys Val Gln His His Met Ala	
	1830 1835 1840
CCC ACT AAA GAT GAG TTT GAC TGC AAA GCC TGG GCT TAT TTC TCT GAT	6582
Pro Thr Lys Asp Glu Phe Asp Cys Lys Ala Trp Ala Tyr Phe Ser Asp	
	1845 1850 1855
GTT GAC CTG GAA AAA GAT GTG CAC TCA GGC CTG ATT GGA CCC CTT CTG	6630
Val Asp Leu Glu Lys Asp Val His Ser Gly Leu Ile Gly Pro Leu Leu	
	1860 1865 1870 1875
GTC TGC CAC ACT AAC ACA CTG AAC CCT GCT CAT GGG AGA CAA GTG ACA	6678
Val Cys His Thr Asn Thr Leu Asn Pro Ala His Gly Arg Gln Val Thr	
	1880 1885 1890
GTA CAG GAA TTT GCT CTG TTT TTC ACC ATC TTT GAT GAG ACC AAA AGC	6726
Val Gln Glu Phe Ala Leu Phe Phe Thr Ile Phe Asp Glu Thr Lys Ser	
	1895 1900 1905
TGG TAC TTC ACT GAA AAT ATG GAA AGA AAC TGC AGG GCT CCC TGC AAT	6774
Trp Tyr Phe Thr Glu Asn Met Glu Arg Asn Cys Arg Ala Pro Cys Asn	
	1910 1915 1920
ATC CAG ATG GAA GAT CCC ACT TTT AAA GAG AAT TAT CGC TTC CAT GCA	6822
Ile Gln Met Glu Asp Pro Thr Phe Lys Glu Asn Tyr Arg Phe His Ala	
	1925 1930 1935
ATC AAT GGC TAC ATA ATG GAT ACA CTA CCT GGC TTA GTA ATG GCT CAG	6870
Ile Asn Gly Tyr Ile Met Asp Thr Leu Pro Gly Leu Val Met Ala Gln	
	1940 1945 1950 1955
GAT CAA AGG ATT CGA TGG TAT CTG CTC AGC ATG GGC AGC AAT GAA AAC	6918
Asp Gln Arg Ile Arg Trp Tyr Leu Leu Ser Met Gly Ser Asn Glu Asn	
	1960 1965 1970
ATC CAT TCT ATT CAT TTC AGT GGA CAT GTG TTC ACT GTA CGA AAA AAA	6966
Ile His Ser Ile His Phe Ser Gly His Val Phe Thr Val Arg Lys Lys	
	1975 1980 1985
GAG GAG TAT AAA ATG GCA CTG TAC AAT CTC TAT CCA GGT GTT TTT GAG	7014
Glu Glu Tyr Lys Met Ala Leu Tyr Asn Leu Tyr Pro Gly Val Phe Glu	

1990	1995	2000	
ACA GTG GAA ATG TTA CCA TCC AAA GCT GGA ATT TGG CGG GTG GAA TGC Thr Val Glu Met Leu Pro Ser Lys Ala Gly Ile Trp Arg Val Glu Cys 2005 2010 2015			7062
CTT ATT GGC GAG CAT CTA CAT GCT GGG ATG AGC ACA CTT TTT CTG GTG Leu Ile Gly Glu His Leu His Ala Gly Met Ser Thr Leu Phe Leu Val 2020 2025 2030 2035			7110
TAC AGC AAT AAG TGT CAG ACT CCC CTG GGA ATG GCT TCT GGA CAC ATT Tyr Ser Asn Lys Cys Gln Thr Pro Leu Gly Met Ala Ser Gly His Ile 2040 2045 2050			7158
AGA GAT TTT CAG ATT ACA GCT TCA GGA CAA TAT GGA CAG TGG GCC CCA Arg Asp Phe Gln Ile Thr Ala Ser Gly Gln Tyr Gly Gln Trp Ala Pro 2055 2060 2065			7206
AAG CTG GCC AGA CTT CAT TAT TCC GGA TCA ATC AAT GCC TGG AGC ACC Lys Leu Ala Arg Leu His Tyr Ser Gly Ser Ile Asn Ala Trp Ser Thr 2070 2075 2080			7254
AAG GAG CCC TTT TCT TGG ATC AAG GTG GAT CTG TTG GCA CCA ATG ATT Lys Glu Pro Phe Ser Trp Ile Lys Val Asp Leu Leu Ala Pro Met Ile 2085 2090 2095			7302
ATT CAC GGC ATC AAG ACC CAG GGT GCC CGT CAG AAG TTC TCC AGC CTC Ile His Gly Ile Lys Thr Gln Gly Ala Arg Gln Lys Phe Ser Ser Leu 2100 2105 2110 2115			7350
TAC ATC TCT CAG TTT ATC ATC ATG TAT AGT CTT GAT GGG AAG AAG TGG Tyr Ile Ser Gln Phe Ile Ile Met Tyr Ser Leu Asp Gly Lys Lys Trp 2120 2125 2130			7398
CAG ACT TAT CGA GGA AAT TCC ACT GGA ACC TTA ATG GTC TTC TTT GGC Gln Thr Tyr Arg Gly Asn Ser Thr Gly Thr Leu Met Val Phe Phe Gly 2135 2140 2145			7446
AAT GTG GAT TCA TCT GGG ATA AAA CAC AAT ATT TTT AAC CCT CCA ATT Asn Val Asp Ser Ser Gly Ile Lys His Asn Ile Phe Asn Pro Pro Ile 2150 2155 2160			7494
ATT GCT CGA TAC ATC CGT TTG CAC CCA ACT CAT TAT AGC ATT CGC AGC Ile Ala Arg Tyr Ile Arg Leu His Pro Thr His Tyr Ser Ile Arg Ser 2165 2170 2175			7542
ACT CTT CGC ATG GAG TTG ATG GGC TGT GAT TTA AAT AGT TGC AGC ATG Thr Leu Arg Met Glu Leu Met Gly Cys Asp Leu Asn Ser Cys Ser Met 2180 2185 2190 2195			7590
CCA TTG GGA ATG GAG AGT AAA GCA ATA TCA GAT GCA CAG ATT ACT GCT Pro Leu Gly Met Glu Ser Lys Ala Ile Ser Asp Ala Gln Ile Thr Ala 2200 2205 2210			7638
TCA TCC TAC TTT ACC AAT ATG TTT GCC ACC TGG TCT CCT TCA AAA GCT Ser Ser Tyr Phe Thr Asn Met Phe Ala Thr Trp Ser Pro Ser Lys Ala 2215 2220 2225			7686

CGA CTT CAC CTC CAA GGG AGG AGT AAT GCC TGG AGA CCT CAG GTG AAT	7734
Arg Leu His Leu Gln Gly Arg Ser Asn Ala Trp Arg Pro Gln Val Asn	
2230 2235 2240	
AAT CCA AAA GAG TGG CTG CAA GTG GAC TTC CAG AAG ACA ATG AAA GTC	7782
Asn Pro Lys Glu Trp Leu Gln Val Asp Phe Gln Lys Thr Met Lys Val	
2245 2250 2255	
ACA GGA GTA ACT ACT CAG GGA GTA AAA TCT CTG CTT ACC AGC ATG TAT	7830
Thr Gly Val Thr Thr Gln Gly Val Lys Ser Leu Leu Thr Ser Met Tyr	
2260 2265 2270 2275	
GTG AAG GAG TTC CTC ATC TCC AGC AGT CAA GAT GGC CAT CAG TGG ACT	7878
Val Lys Glu Phe Leu Ile Ser Ser Ser Gln Asp Gly His Gln Trp Thr	
2280 2285 2290	
CTC TTT TTT CAG AAT GGC AAA GTA AAG GTT TTT CAG GGA AAT CAA GAC	7926
Leu Phe Phe Gln Asn Gly Lys Val Lys Val Phe Gln Gly Asn Gln Asp	
2295 2300 2305	
TCC TTC ACA CCT GTG GTG AAC TCT CTA GAC CCA CCG TTA CTG ACT CGC	7974
Ser Phe Thr Pro Val Val Asn Ser Leu Asp Pro Pro Leu Leu Thr Arg	
2310 2315 2320	
TAC CTT CGA ATT CAC CCC CAG AGT TGG GTG CAC CAG ATT GCC CTG AGG	8022
Tyr Leu Arg Ile His Pro Gln Ser Trp Val His Gln Ile Ala Leu Arg	
2325 2330 2335	
ATG GAG GTT CTG GGC TGC GAG GCA CAG GAC CTC TAC TGAGGGTGGC	8068
Met Glu Val Leu Gly Cys Glu Ala Gln Asp Leu Tyr	
2340 2345 2350	
CACTGCAGCA CCTGCCACTG CCGTCACCTC TCCCTCCTCA GCTCCAGGGC AGTGTCCCTC	8128
CCTGGCTTGC CTTCTACCTT TGTGCTAAAT CCTAGCAGAC ACTGCCTTGA AGCCTCCTGA	8188
ATTAACATATC ATCAGTCCTG CATTTCTTTG GTGGGGGGCC AGGAGGGTGC ATCCAATTTA	8248
ACTTAACTCT TACCTATTTT CTGCAGCTGC TCCCAGATTA CTCCTTCCTT CCAATATAAC	8308
TAGGCAAAAA GAAGTGAGGA GAAACCTGCA TGAAAGCATT CTTCCCTGAA AAGTTAGGCC	8368
TCTCAGAGTC ACCACTTCCT CTGTTGTAGA AAACTATGT GATGAACTT TGAAAAAGAT	8428
ATTTATGATG TTAACCTGTT TATTGCAGCT TATAATGGTT ACAAATAAAG CAATAGCATC	8488
ACAAATTTCA CAAATAAAGC ATTTTTCCTCA CTGCATTCTA GTTGTGGTTT GTCCAAACTC	8548
ATCAATGTAT CTTATCATGT CTGGATCCCC GGGTGGCATC CCTGTGACCC CTCCCCAGTG	8608
CCTCTCCTGG CCCTGGAAGT TGCCACTCCA GTGCCACCA GCCTTGTCTT AATAAAATTA	8668
AGTTGCATCA TTTTGTCTGA CTAGGTGTCC TTCTATAATA TTATGGGGTG GAGGGGGGTG	8728
GTATGGAGCA AGGGGCAAGT TGGGAAGACA ACCTGTAGGG CCTGCGGGGT CTATTTCGGGA	8788

ACCAAGCTGG	AGTGCAGTGG	CACAATCTTG	GCTCACTGCA	ATCTCCGCCT	CCTGGGTTCA	8848
AGCGATTCTC	CTGCCTCAGC	CTCCCGAGTT	GTTGGGATTC	CAGGCATGCA	TGACCAGGCT	8908
CAGCTAATTT	TTGTTTTTTT	GGTAGAGACG	GGGTTTCACC	ATATTGGCCA	GGCTGGTCTC	8968
CAACTCCTAA	TCTCAGGTGA	TCTACCCACC	TTGGCCTCCC	AAATTGCTGG	GATTACAGGC	9028
GTGAACCACT	GCTCCCTTCC	CTGTCCTTCT	GATTTTAAAA	TAACTATACC	AGCAGGAGGA	9088
CGTCCAGACA	CAGCATAGGC	TACCTGCCAT	GCCCAACCGG	TGGGACATTT	GAGTTGCTTG	9148
CTTGGCACTG	TCCTCTCATG	CGTTGGGTCC	ACTCAGTAGA	TGCCTGTTGA	ATTCGTAATC	9208
ATGGTCATAG	CTGTTTCCTG	TGTGAAATTG	TTATCCGCTC	ACAATTCCAC	ACAACATACG	9268
AGCCGGAAGC	ATAAAGTGTA	AAGCCTGGGG	TGCCTAATGA	GTGAGCTAAC	TCACATTAAT	9328
TGCGTTGCGC	TCACTGCCCC	CTTTCCAGTC	GGGAAACCTG	TCGTGCCAGC	TGCATTAATG	9388
AATCGGCCAA	CGCGCGGGGA	GAGGCGGTTT	GCGTATTGGG	CGCTCTTCCG	CTTCCTCGCT	9448
CACTGACTCG	CTGCGCTCGG	TCGTTCCGGCT	GCGGCGAGCG	GTATCAGCTC	ACTCAAAGGC	9508
GGTAATACGG	TTATCCACAG	AATCAGGGGA	TAACGCAGGA	AAGAACATGT	GAGCAAAAAG	9568
CCAGCAAAAG	GCCAGGAACC	GTAAAAAGGC	CGCGTTGCTG	GCGTTTTTCC	ATAGGCTCCG	9628
CCCCCTGAC	GAGCATCACA	AAAATCGACG	CTCAAGTCAG	AGGTGGCGAA	ACCCGACAGG	9688
ACTATAAAGA	TACCAGGCGT	TTCCCCCTGG	AAGCTCCCTC	GTGCGCTCTC	CTGTTCCGAC	9748
CCTGCCGCTT	ACCGGATACC	TGTCCGCCTT	TCTCCCTTCG	GGAAGCGTGG	CGCTTTCTCA	9808
TAGCTCACGC	TGTAGGTATC	TCAGTTCGGT	GTAGGTCGTT	CGCTCCAAGC	TGGGCTGTGT	9868
GCACGAACCC	CCCGTTCAGC	CCGACCGCTG	CGCCTTATCC	GGTAACTATC	GTCTTGAGTC	9928
CAACCCGGTA	AGACACGACT	TATCGCCACT	GGCAGCAGCC	ACTGGTAACA	GGATTAGCAG	9988
AGCGAGGTAT	GTAGGCGGTG	CTACAGAGTT	CTTGAAGTGG	TGGCCTAACT	ACGGCTACAC	10048
TAGAAGGACA	GTATTTGGTA	TCTGCGCTCT	GCTGAAGCCA	GTTACCTTCG	GAAAAAGAGT	10108
TGGTAGCTCT	TGATCCGGCA	AACAAACCAC	CGCTGGTAGC	GGTGGTTTTT	TTGTTTGCAA	10168
GCAGCAGATT	ACGCGCAGAA	AAAAAGGATC	TCAAGAAGAT	CCTTTGATCT	TTTCTACGGG	10228
GTCTGACGCT	CAGTGGAACG	AAAACTCACG	TTAAGGGATT	TTGGTCATGA	GATTATCAAA	10288
AAGGATCTTC	ACCTAGATCC	TTTTAAATTA	AAAATGAAGT	TTTAAATCAA	TCTAAAGTAT	10348
ATATGAGTAA	ACTTGGTCTG	ACAGTTACCA	ATGCTTAATC	AGTGAGGCAC	CTATCTCAGC	10408
GATCTGTCTA	TTTCGTTTCT	CCATAGTTGC	CTGACTCCCC	GTCGTGTAGA	TAACTACGAT	10468
ACGGGAGGGC	TTACCATCTG	GCCCCAGTGC	TGCAATGATA	CCGCGAGACC	CACGCTCACC	10528

GGCTCCAGAT TTATCAGCAA TAAACCAGCC AGCCGGAAGG GCCGAGCGCA GAAGTGGTCC 10588
TGCAACTTTA TCCGCCTCCA TCCAGTCTAT TAATTGTTGC CGGGAAGCTA GAGTAAGTAG 10648
TTCGCCAGTT AATAGTTTGC GCAACGTTGT TGCCATTGCT ACAGGCATCG TGGTGTACAG 10708
CTCGTCGTTT GGTATGGCTT CATTAGCTC CGGTTCCCAA CGATCAAGGC GAGTTACATG 10768
ATCCCCCATG TTGTGCAAAA AAGCGGTTAG CTCCTTCGGT CCTCCGATCG TTGTCAGAAG 10828
TAAGTTGGCC GCAGTGTAT CACTCATGGT TATGGCAGCA CTGCATAATT CTCTTACTGT 10888
CATGCCATCC GTAAGATGCT TTTCTGTGAC TGGTGAGTAC TCAACCAAGT CATTCTGAGA 10948
ATAGTGTATG CGGCGACCGA GTTGCTCTTG CCCGGCGTCA ATACGGGATA ATACCGCGCC 11008
ACATAGCAGA ACTTTAAAAG TGCTCATCAT TGGAAAACGT TCTTCGGGGC GAAACTCTC 11068
AAGGATCTTA CCGCTGTTGA GATCCAGTTC GATGTAACCC ACTCGTGCAC CCAACTGATC 11128
TTCAGCATCT TTTACTTTCA CCAGCGTTTC TGGGTGAGCA AAAACAGGAA GGCAAAATGC 11188
CGCAAAAAG GGAATAAGGG CGACACGGAA ATGTTGAATA CTCATACTCT TCCTTTTTCA 11248
ATATTATTGA AGCATTTATC AGGGTTATTG TCTCATGAGC GGATACATAT TTGAATGTAT 11308
TTAGAAAAAT AAACAAATAG GGGTTCGCG CACATTTCCC CGAAAAGTGC CACCTGACGT 11368
CTAAGAAACC ATTATTATCA TGACATTAAC CTATAAAAAT AGGCGTATCA CGAGGCCCTT 11428
TCGTCTCGCG CGTTTCGGTG ATGACGGTGA AAACCTCTGA CACATGCAGC TCCCGGAGAC 11488
GGTCACAGCT TGTCTGTAAG CGGATGCCGG GAGCAGACAA GCCCGTCAGG GCGCGTCAGC 11548
GGGTGTTGGC GGGTGTGGG GCTGGCTTAA CTATGCGGCA TCAGAGCAGA TTGTACTGAG 11608
AGTGCACCAT ATGCGGTGTG AAATACCGCA CAGATGCGTA AGGAGAAAAT ACCGCATCAG 11668
GCGCCATTCT CCATTCAGGC TCGCCACTG TTGGGAAGGG CGATCGGTGC GGGCCTCTTC 11728
GCTATTACGC CAGCTGGCGA AAGGGGGATG TGCTGCAAGG CGATTAAGTT GGGTAACGCC 11788
AGGGTTTTCC CAGTCACGAC GTTGTA AAC GACGGCCAGT GCCAAGCTTG GGCTGCAG 11846

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 211 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

ATTGAACCAA GAAGCTTCTC CCAGGTAAGT TGCTAATAAA GCTTGGCAAG AGTATTTCAA 60
 GGAAGATGAA GTCATTAAC TATGCAAAATG CTTCTCAGGC ACCTAGGAAA ATGAGGATGT 120
 GAGGCATTTC TACCCACTTG GTACATAAAA TTATTGCTTT TCCTCTTCTT TTTTCTCCA 180
 GAACCCACCA GTCTTGAAAC GCCATCAACG G 211

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 126 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GTTGGTATCC TTTTACAGC ACAACTTAAT GAGACAGATA GAAACTGGTC TTGTAGAAAC 60
 AGAGTAGTCG CCTGCTTTTC TGCCAGGTGC TGA CTCTCTCT CCCCTGGGCT GTTTTCATTT 120
 TCTCAG 126

(2) INFORMATION FOR SEQ ID NO:7:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 126 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GTAAGTATCC TTTTACAGC ACAACTTAAT GAGACAGATA GAAACTGGTC TTGTAGAAAC 60
 AGAGTAGTCG CCTGCTTTTC TGCCAGGTGC TGA CTCTCTCT CCCCTTCTCT TTTTCTCTT 120
 TCTCAG 126

(2) INFORMATION FOR SEQ ID NO:8:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 10 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GCCACCAUGG 10

(2) INFORMATION FOR SEQ ID NO:9:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 100 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

```
AGGTTAATTT TTAAAAAGCA GTCAAAGTC CAAGTGGCCC TTGCGAGCAT TTACTCTCTC    60
TGTTTGCTCT GGTTAATAAT CTCAGGAGCA CAAACATTCC                            100
```

(2) INFORMATION FOR SEQ ID NO:10:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 223 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

```
CTTTCTCTTT TCTTTTACAT GAAGGGTCTG GCAGCCAAAG CAATCACTCA AAGTTCAAAC    60
CTTATCATTT TTTGCTTTGT TCCTCTTGGC CTTGGTTTTG TACATCAGCT TTGAAAATAC    120
CATCCCAGGG TTAATGCTGG GGTAAATTTA TAACTAAGAG TGCTCTAGTT TTGCAATACA    180
GGACATGCTA TAAAAATGGA AAGATGTTGC TTTCTGAGAG ATA                        223
```

(2) INFORMATION FOR SEQ ID NO:11:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 90 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

```
AGAUCUCGAG AAAGCUAACA ACAAAGAACA ACAAACAACA AUCAGGAUAA CAAGAACGAA    60
ACAAUAACAG CCACCAUGGA AAUAGAGCUC                                    90
```